



Escuela
Politécnica
Superior

Retroalimentación integrada centrada en el usuario para la toma de decisiones



Grado en Ingeniería Informática

Trabajo Fin de Grado

Autor:

Paulino Corchero León

Tutor/es:

Juan Carlos Trujillo Mondéjar

Alejandro Mate Morga



Universitat d'Alacant
Universidad de Alicante

Índice

Prefacio	5
1. Introducción	7
2. Fases del Proyecto.....	15
2.1 Decisiones estratégicas	15
2.2 Diseño.....	22
2.2.1 Almacén de Datos	23
2.2.2 ETL	24
2.2.3 Modelos Analíticos (OLAP)	26
2.2.4 Informes y Cuadros de Mando	29
3. Retroalimentación integrada centrada en el usuario para la toma de decisiones	33
4. Herramientas.....	38
4.1 Almacén de Datos, ETL y sistema OLAP	38
4.1.1 Cuadrante Mágico de Gartner para plataformas de BI y Analíticas 2015	38
4.1.2 Herramienta seleccionada	40
4.2 Herramientas de procesamiento de modelos OLAP	41
4.2.1 Data Discovery	43
4.2.2 Herramienta seleccionada	49
5. Caso Práctico	50
5.1 Almacén de Datos	54
5.1.1 Esquema Aux.....	54
5.1.2 Esquema Stage	55
5.1.3 Esquema DWH.....	57
5.2 Extracción, transformación y carga de datos.....	61
5.2.1 Estructura de BBDD	62
5.2.2 Carga de datos.....	63
5.3 Modelos analíticos (OLAP)	65
5.3.1 Dimensiones	65
5.3.2 Grupos de medidas	79
5.3.3 Esquema estrella	83
5.3.4 Roles seguridad	86
5.4 Informes y cuadros de mando	87
5.4.1 Analistas	88
5.4.2 Informes funcionales.....	88
5.4.3 Cuadro de mando (CMI)	93
5.5 Retroalimentación de Sistema	97
5.5.1 Análisis de la información para la toma de decisiones	98
5.5.2 Identificación y clasificación DAFO.....	98
5.5.3 Toma de Decisiones	99
5.5.4 Fichero de Retroalimentación	100

5.5.5	Análisis de resultados.....	101
6.	Conclusiones	102
7.	Referencias.....	104

Prefacio

Este proyecto viene suscitado por el contenido desarrollado en la asignatura optativa de Inteligencia de Negocio y Gestión de Procesos. Tal asignatura pertenece al 4º Curso del plan de estudios de la titulación de Grado en Ingeniería Informática (plan 2010) de la Universidad de Alicante (UA). Los temas que son tratados en dicho trabajo pertenecen al área de conocimiento de Lenguajes y Sistemas Informáticos (LSI) y el proyecto ha sido realizado bajo la supervisión, seguimiento y evaluación del Catedrático de Universidad D. Juan Carlos Trujillo Mondéjar, dentro del seno del Grupo de Investigación Lucentia que el mismo dirige y de Alejandro Mate Morga Doctor en Informática por la Universidad de Alicante y miembro del Grupo de Investigación Lucentia.

La motivación principal para el desarrollo del presente trabajo ha sido el gran interés que me han suscitado los temas estudiados en la asignatura optativa, que trata fundamentalmente sobre el diseño y explotación de los almacenes de datos, aplicados en entornos corporativos.

El objetivo fundamental del trabajo realizado consiste en completar la estrategia de Business Intelligence o BI [\[1\]](#) dirigida por el usuario o DATA DISCOVERY [\[2\]](#), en el que el usuario descubre los datos para tomar decisiones que repercuten directamente con la estrategia de la empresa.

Para conseguir el objetivo del proyecto se han aplicado técnicas de estrategia empresarial centradas en metodologías bajo metodologías de análisis DAFO [\[3\]](#) (Debilidades, Amenazas, Fortalezas y Oportunidades.), esta metodología de estudio permite evaluar la situación de una empresa o un proyecto, permitiendo analizar las ventas de una empresa buscando las debilidades, fortalezas,

amenazas y oportunidades. Este análisis retro alimentará el sistema de Business Intelligence o BI, permitiendo incorporar la experiencia de la toma de decisiones en el sistema de BI.

La experiencia de la toma de decisiones, permite al usuario validar de manera objetiva el éxito o fracaso de las decisiones adoptadas, ofreciendo al usuario y empresa el conocimiento y experiencia de decisiones de negocio anteriores.

La experiencia de la toma de decisiones queda almacenada, favoreciendo claramente la continuidad de negocio. Las experiencias y la eficiencia estas no se pierden y pasa a formar parte de la base de conocimiento de la empresa.

1. Introducción

Desde tiempos remotos los datos han sido registrados por el hombre en algún tipo de soporte (papel, piedra, madera), con la llegada de la informática en los 50, la los soportes cambiaron, la cantidad de información almacenada y procesada no ha dejado de crecer desde entonces.

En los años 70 Edgar Frank Codd [\[4\]](#), introduce los conceptos de las Base de Datos Relacionales, que se basan en relaciones las cuales se podían considerar en forma lógica como Tuplas, propuestos en "Las doce reglas de Codd". En 1945 se crean los primeros sistemas Planificación de los requerimientos de material (MRP), desarrollados para responder a las necesidades logísticas de ejércitos. A finales de la Segunda Guerra Mundial, el ejército de Estados Unidos empezó a usar programas informáticos para gestionar las complejas tareas de producción y logística del esfuerzo bélico. Aplicarlos a la industria era aún inviable, porque solo las instituciones militares tenían computadoras en esta época. En la década de los 70, la industria se podía permitir acceder a sistemas informático. Los MRP, aplicaciones de planificación militar se consideran el origen de los actuales programas de planificación de recursos empresariales o ERP [\[5\]](#) Enterprise Resource Planning. En esta década también se fundaron la mayoría de los proveedores de ERP que conocemos actualmente, como SAP (1972), J.D. Edwards (1977) o Baan (1978).

Hoy en día los sistemas ERP, también conocidos como sistemas OLTP [\[6\]](#) OnLine Transaction Processin más importantes almacenan la información de manera transaccional. Esta información no se destruye, simplemente se transforma: inventario, transacciones de cliente, transacciones de proveedor, transiciones de actividad registradas en sistemas CRM [\[7\]](#) Customer Relationship Management.

Las transacciones de una empresa permiten por ejemplo, reconstruir inventarios a fecha y saber a ciencia cierta que existencias había de un producto en un momento determinado. Estas transacciones son en sí un registro de lo que ha sucedido en la empresa y permiten a la empresa analizar la información acerca de situaciones pasadas para tomar decisiones estratégicas de cara al futuro.

Estas transacciones permiten las empresas resolver las necesidades funcionales de todas las áreas que comprenden la empresa, almacén, administración, fabricación, comercialización, distribución... Lamentablemente debido al volumen información y la estructura en la que se almacenan las transacciones el análisis de la información en este tipo de sistemas son extremadamente costosos, porque generalmente están normalizadas bajo la 5ª forma normal (mínimo 3ª forma norma), también conocida como forma normal de proyección-uni6n (PJ/NF). La 5ª forma normal es un nivel de normalizaci6n de bases de datos dise1ado para reducir redundancia en las bases de datos relacionales que guardan hechos multi-valores aislando semánticamente relaciones m6ltiples relacionadas. Esta normalizaci6n de los sistemas ERP que evitan la redundancia de los datos, causa que resulte prácticamente impracticable la realizaci6n de análisis complejos sobre las bases de datos.

Un análisis complejo directamente sobre estas bases de datos, requieren un alto conocimiento técnico, debido al elevado número de entidades (tablas), relaciones y su semántica. Por este motivo resulta impracticable la realizaci6n de análisis sobre estas bases de datos. Sirva como ejemplo la siguiente consulta, efectuada sobre el ERP Microsoft Dynamics Ax 2009 que se puede observar en la Figura 1.

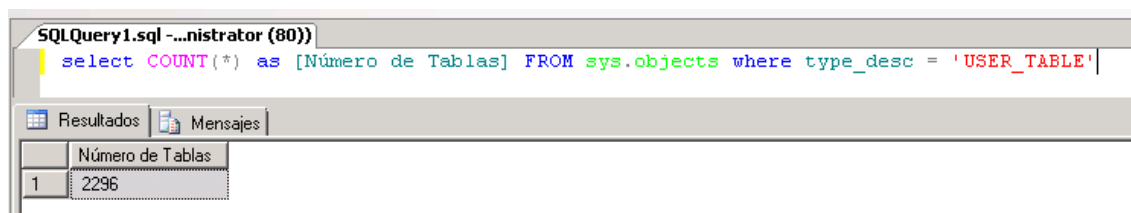


Figura 1: Consulta Tablas Microsoft Dynamics

En definitiva, en el sistema a estudio, Microsoft Dynamics Axapta 2009 sp1 [8], nos encontramos con 2.296 tablas. Entender el significado de estas entidades, sus atributos y sus relaciones, conlleva un esfuerzo insostenible de cara al análisis.

El análisis de la información es un tema generalmente complejo, ya que la cantidad de información suele ser impracticable para sistemas de reporting clásicos, donde el entorno es prácticamente estático y basan el análisis en informes imprimibles, que por norma general no tienen capacidades de síntesis de la propia información.

La forma normal, de este tipo base de datos, reducen el redundancia de datos, conllevan que las consultas necesarias para realizar un análisis complejo, requieran mucho conocimiento técnico y generalmente al estar tan distribuida la información el rendimiento de la propia consulta comprometa el rendimiento global del sistema.

El Reporting generado desde estos sistemas, permiten un filtrado muy potente, permitiendo al usuario en tiempo de ejecución, enlazar entidades adyacentes a las entidades implementadas en el propio informe. Siendo a priori positivo, puede generar ineficiencias en el rendimiento global de la aplicación, ya que por cada entidad que se añade a la consulta el sistema realiza los join o enlaces necesarios para acceder a las entidades. Estos enlaces pueden literalmente llegar a bloquear el servidor de base de datos. En conclusión el rendimiento del sistema baja drásticamente cuando se usa reporting analítico, ya

que principalmente la 5 forma normal conlleva un elevado número de entidades enlazadas, este tipo de instrucciones (JOIN y GROUP BY principalmente), afectan de manera directa al rendimiento del servidor de base de datos

Desde el punto de partida de un ERP como fuente única de análisis de datos nos encontramos, que este tipo de sistemas están pensados para optimizar el espacio, favoreciendo la inserción rápida de registros. William Kent [\[9\]](#) uno de los fundadores y primer presidente del comité ANSI X3H7 en Gestión de la Información de objetos, redactó un memorable resumen de la definición de Codd de la 5ª forma normal, en el documento autodenominado “A Simple Guide to Five Normal Forms in Relational Database Theory” [\[10\]](#), en el que el lector puede entender las formas normales definidas en la teoría de bases de datos relacionales representan pautas para el diseño de registro. Bajo este diseño orientado a la optimización del espacio los ERP actuales consiguen que las empresas almacenen un gran nivel de detalle de las transacciones en cualquier área, que permiten responder a las necesidades funcionales de todas estas áreas. Todas las transacciones dentro de estos ERP tienen una orientación funcional, una factura, un inventario disponible, la declaración de un impuesto...

La orientación funcional de las áreas de una empresa implican normalmente desarrollos ad-hoc [\[11\]](#), una solución específicamente elaborada para un problema o fin preciso y, por tanto, no generalizable ni utilizable para otros propósitos. Debido al uso de la 5 forma normal en los ERP, el desarrollador debe conocer las entidades con las que opera, generalmente un informe simple hace uso de al menos 5 entidades, el desarrollo en este tipo de sistemas es prácticamente estático, una vez se adquiere la comprensión de las entidades necesarias, el informe en si refleja unidades de medida agrupadas por estas entidades. Estas entidades agrupadas o las unidades de medida no pueden ser modificadas en tiempo de ejecución, y cualquier variación ha de ser implementada por un desarrollador. En conclusión el coste de implementación en estos sistemas incurre en costes elevados.

La conclusión es que el diseño de los ERP, ofrece principalmente respuestas funcionales, el uso de estos sistemas para dar respuestas analíticas, tiene un coste muy elevado de implementación y afectan de manera directa al rendimiento del propio sistema, cuando crece el volumen de información y los requerimientos del análisis, el desarrollador tiene que implementar sus informes pensando cada vez más en el rendimiento, y esto genera que los desarrollos de demoren creando un cuello de botella en el servicio de la información, lo cual se traduce por norma general en que la información no llega a tiempo.

Estratificar el modelo de datos des normalizando, optimiza el desempeño de la base de datos por medio de agregar datos redundantes, permitiendo que los datos se reagrupen generando navegaciones verticales desde datos generales a datos particulares, esto facilita análisis más fluidos, mejorar significativamente el rendimiento de las consultas y reducir los conocimientos técnico necesarios. Edgar Frank Codd, también acuñó el término OLAP y redactó las doce leyes del procesamiento analítico informático. En la Figura 2 se pueden observar las principales características de un procesamiento OLTP vs OLAP

OLTP	OLAP
Diseñados para permitir una inserción rápida de datos (registros)	Diseñados para permitir consultas de alto rendimiento a grandes velocidades
Permiten almacenar transacciones al mayor nivel de detalle	Facilitan la consolidación de los datos de manera automática y óptima
Reducen el espacio de almacenamiento mediante la aplicación de Formas Normales al diseño, evitando, entre otras cosas, la redundancia de datos	Utilizan espacio para almacenar las estructuras que faciliten la rápida respuesta de las consultas
Generalmente utilizan Tablas relacionadas para almacenar los datos	Generalmente utilizan Cubos, formados de Dimensiones

Figura 2: Comparativa OLTP vs OLAP

Todas estas razones derivan y motivan, la búsqueda de un sistema que permita hacer frente al problema analítico que surgen en todas las organizaciones.

La arquitectura de DataWareHouse [\[12\]](#) de datos, nació en la década de 1980 como un modelo arquitectónico diseñado para apoyar el flujo de datos de los sistemas de BI. En un almacén de datos lo que se quiere es contener datos que son necesarios o útiles para una organización, es decir, que se utiliza como un repositorio de datos, para posteriormente transformarlos en información útil para el usuario.

Según William H. Inmon [\[13\]](#), reconocido por muchos como el padre de los almacenes de datos, un almacén de datos es:

“Una colección de datos orientados por tema, integrados, variables en el tiempo y no volátiles que se emplea como apoyo a la toma de decisiones estratégicas”

El autor principal de uno de los libros más vendidos “The Data Warehouse Toolkit”, Ralph Kimball [\[14\]](#), define un almacén de datos como: "una copia de las transacciones de datos específicamente estructurada para la consulta y el análisis".

Esta arquitectura puede plantearse desde dos puntos de vista, arquitectura dirigida por las transacciones (BI clásico [\[15\]](#)) o arquitecturas dirigidas por el usuario (Data Discovery).

A medida que más organizaciones recurren a análisis de datos para ofrecer valor, estas dos tendencias de BI se disputan la mayor parte de las dichas organizaciones.

El BI clásico ha tenido durante mucho tiempo la reputación de ser difícil de implementar e igualmente difícil de usar, orientados principalmente a facilitar el proceso de toma de decisiones de alto nivel. Partiendo de esta percepción, herramientas de descubrimiento de datos nacieron y comenzaron a ganar fuerza. En muchos sentidos, el descubrimiento de datos respondió a la necesidad de rapidez en la empresa de hoy.

Las soluciones tradicionales de BI basan su objetivo de diseño en entregar todos los temas de análisis posible y una imagen universal de la empresa, mientras que las herramientas de descubrimiento de datos brotaron para resolver la necesidad de velocidad.

Ambos sistemas reducen los costes de reporting que nada tienen que ver con los costes de implementación de informes dentro de un ERP transaccional, y optar por la opción de un sistema dirigido por el usuario, favorece claramente la necesidad de rapidez en la empresas, permitiendo que la información no se busque, se descubra, dando sentido al termino DATA DISCOVERY.

El rendimiento que otorga el diseño de estos sistemas facilita la gestión de grandes volúmenes de datos, siendo responsable directo de un rendimiento más eficiente de cara a la explotación de la información. Tanto por diseño como por especialización, el sistema ERP queda prácticamente aislado del sistema de BI, generando dos entornos claramente definidos por sus objetivos. ERP-Funcional y BI-Analítico.

En este punto, el usuario dispone de una arquitectura completa para la de decisiones, las cuales derivan directamente del estudio de los datos facilitados por la arquitectura. Completada la arquitectura, el objetivo del proyecto determina que las decisiones, sus motivaciones (los análisis), acciones (derivadas de las decisiones) y objetivos de estas retro alimenten el sistema, permitiendo que los usuarios validen el éxito o fracaso del proceso de toma de decisiones. El almacenamiento de estas nuevas transacciones permitirá a la empresa generar una base de conocimiento del proceso de toma de decisiones de la propia compañía.

2. Fases del Proyecto

Los sistemas de BI, de forma natural, no incluyen retroalimentación en el sistema. Es por ello que, para llevar a cabo el objetivo último del proyecto, la retroalimentación del sistema centrada en la toma de decisiones, resulta necesario estructurar el proyecto completo desde el principio. Para ello, en este trabajo se toman decisiones de implementación que facilitan la consecución de dicho objetivo. Sin bien el análisis de decisiones se puede implementar directamente sin la necesidad de un sistema de BI, contrastar las decisiones tomadas con el sistema de BI, verifica el resultado de las decisiones con datos reales. Esta retroalimentación permitirá a la empresa detectar rápidamente si los objetivos de las decisiones se han cumplido y generar una base de conocimiento que ayudará a no tomar reiteradamente decisiones erróneas y a identificar las decisiones acertadas.

Una decisión errónea en un contexto, puede ser brillante en otro, es decir no hay decisiones erróneas, por norma general las decisiones erróneas se deben a la falta de información. Entender por qué ha fracasado o triunfado una decisión, forma parte de proceso de aprendizaje, la retroalimentación del sistema tiene como objetivo este aprendizaje.

2.1 Decisiones estratégicas

La primera pregunta que surge a la hora de implementar un sistema de BI tiene que ver con la estrategia, actualmente hay dos estrategias en mercado bien definidas y maduras, el BI clásico y el Data Discovery o tecnologías en memoria (in memory), cada una de ellas aporta características que responden a necesidades concretas.

El sistema de BI Tradicional generalmente es dirigido por analistas e ingenieros informáticos, su diseño [\[16\]](#) e implementación no busca exclusivamente el análisis de la información. La implementación de estos sistemas se ha normalizado y tiene una estructura muy concreta, Data Warehouse, ETL [\[17\]](#) Extract-Transform-Load, OLAP [\[18\]](#) On-Line Analytical Processing y finalmente Servicios de Reporte. El rol que dirige esta solución tiene por norma general altos conocimientos técnicos y poco conocimiento de negocio.

El Data Discovery o tecnologías en memoria (in memory) generalmente es dirigido por consultores o usuarios de negocio, estos buscan responder de manera analítica al negocio, los detalles técnicos por norma general son tratados como problemas menores. El rol que dirige esta solución tiene por norma general altos conocimientos del negocio y poco conocimiento técnico.

Desde esta perspectiva, existen varias alternativas:

1. El Coordinador del proyecto tiene un alto conocimiento del negocio y técnico
2. El Coordinador del proyecto tiene un alto conocimiento del negocio y bajo técnico
3. El Coordinador del proyecto tiene un bajo conocimiento del negocio y alto técnico

Evidentemente el todo el mundo elegiría la primera alternativa, pero siendo realista este coordinador tiene un coste más elevado y es complicado de encontrar un coordinador con alto conocimiento en ambas vertientes.

Entre las alternativas aplicadas al modelo de BI existente existe una cuarta posibilidad que implica la generación de un modelo híbrido de las dos estrategias de mercado planteadas, permitiendo dirigir el proyecto en base al rol necesario en cada momento

Usando como base el modelo de BI tradicional se pueden identificar claramente 4 fases y asignarlas al rol correspondiente.

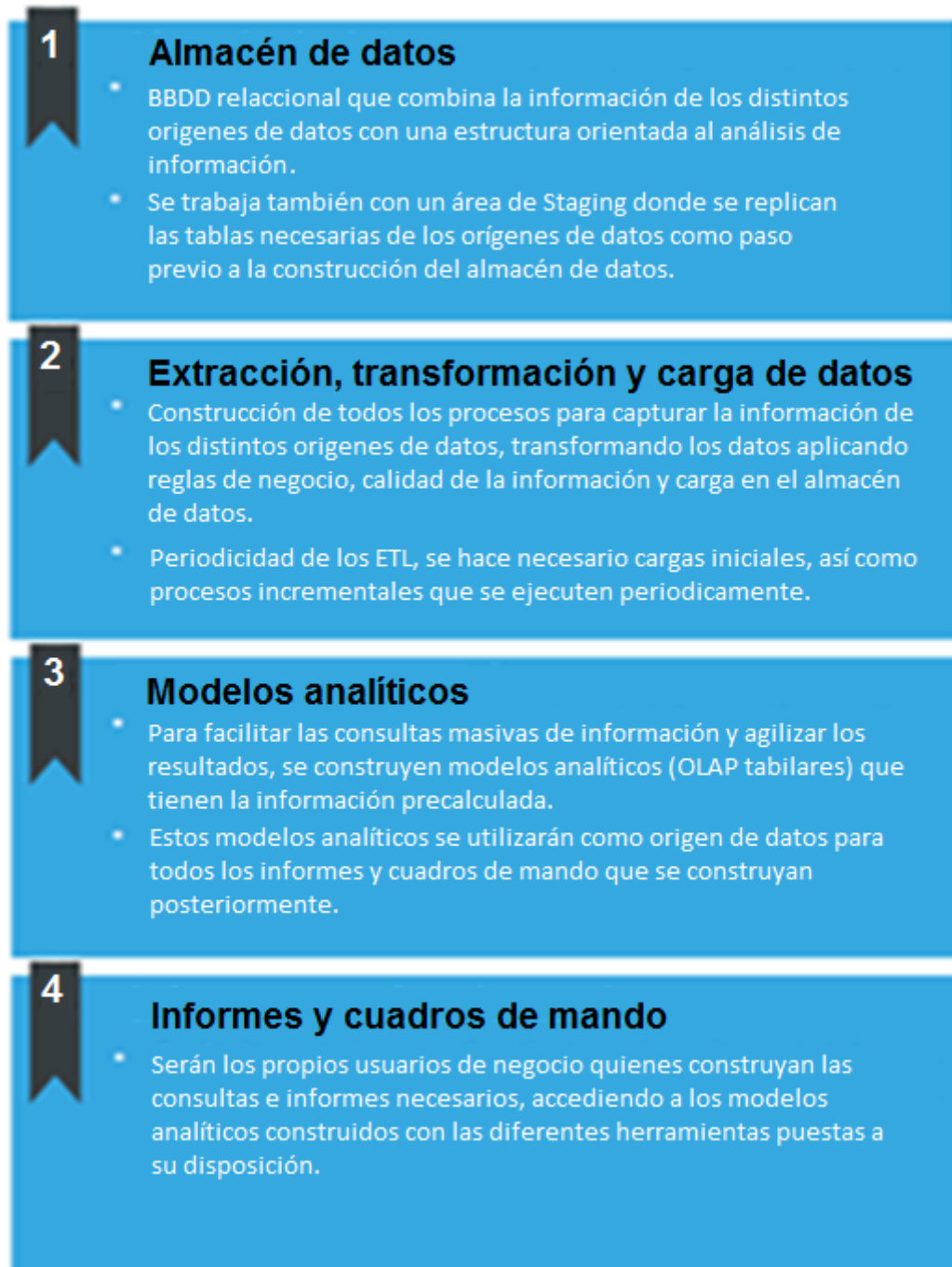


Figura 3: Fases del Proyecto

Usando como base las cuatro fases de un diseño de BI tradicional, la fase final será implementada usando herramientas de Data Discovery, permitiendo que cada especialista tenga su espacio delimitado en responsabilidad y tiempo. Si bien en todas las fases intervienen ambos, las decisiones de cada fase son tomadas por el rol predominante.

Estos Roles centrarán su atención en estrategias distintas:

Especialista TI

1. Rendimiento: El sistema de BI tiene que funcionar de manera independiente al sistema transaccional, si bien obtiene los datos de este, no puede afectar al rendimiento funcional de este (sistemas independientes, ETL controlados, hardware y software correctamente dimensionados)
2. Fiabilidad: El sistema de BI ha de convertirse en un entorno de alta disponibilidad, garantizando la integridad del almacén de datos, los ETL y la disponibilidad de los cubos generados en el modelo analítico.
3. Veracidad: El sistema de BI tiene que ofrecer datos correctos
4. Seguridad: El acceso a los datos tiene que ser seguro, cada rol de usuario, solo puede acceder a estos bajo la directrices de seguridad de la empresa.

Especialista Negocio

1. Facilidad de uso: Orientando al especialista de TI para que el modelo analítico resultante sea de fácil comprensión para cualquier usuario de negocio
2. Formación continua: El sistema debe ser interiorizado a partir de la formación y su uso continuo
3. Reporting: Buscar de manera inequívoca la información relevante para el negocio a partir de:

- a. informes funcionales concretos
- b. CMI / Balanced Scorecard/ DashBoard [\[19\]](#)
- c. Análisis DAFO

Antes de la ejecución del proyecto hay que tomar como referencia el plan estratégico de la empresa, este es representa un programa de actuación que explica claramente que pretende conseguir en la empresa y cómo la empresa propone conseguirlo. Esta programación se plasma en un documento de consenso donde concretamos las grandes decisiones que van a orientar nuestra marcha hacia la gestión excelente. Hay que verificar que en todas las fases se alinean con dicho plan, y que el sistema BI, tiene de tener capacidades de adaptación y cambio, principalmente porque los planes estratégicos pueden variar con el tiempo y el sistema de BI tiene que estar preparado para adaptarse al cambio.

Finalmente el modelo de BI seguirá la siguiente estructura:

1. Almacén de Datos Rol predominante Especialista TI
2. ETL (Extracción, transformación y carga de datos) Rol predominante Especialista TI
3. Modelos analíticos Rol predominante Especialista TI
4. Informes y cuadro de mandos Rol predominante Especialista de Negocio

El proyecto incorpora una nueva fase que lo completa e incorpora un nexo de unión entre el análisis efectuado en punto cuatro y las decisiones aplicadas en base a dicho análisis

5. Retroalimentación integrada centrada en el usuario para la toma de decisiones

Esta estrategia de diseño aporta especialización necesaria en cada fase y facilita la consecución del objetivo de retroalimentación del sistema. En la Figura 4 se

puede observar cómo se rompe la barrera entre el BI tradicional y el Data Discovery, y como estos datos vuelven a incidir sobre el sistema de BI.

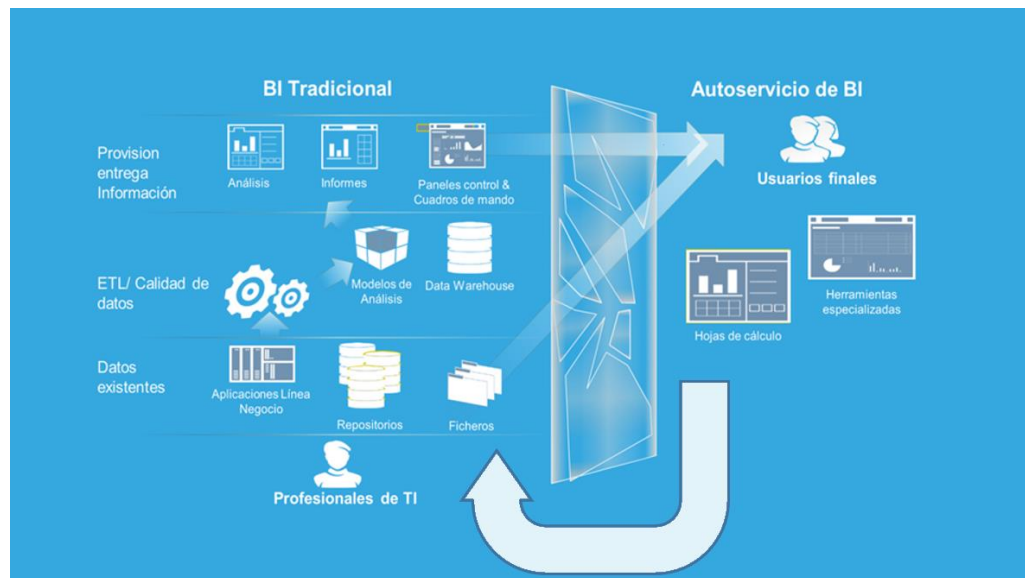


Figura 4: Flujo de datos propuesto

Este diseño usa la siguiente secuencia, que se puede observar en la Figura 5 el sistema queda completo, pasamos de un modelo secuencial compuesto por cuatro fases, a un modelo cíclico compuesto por cinco fases.

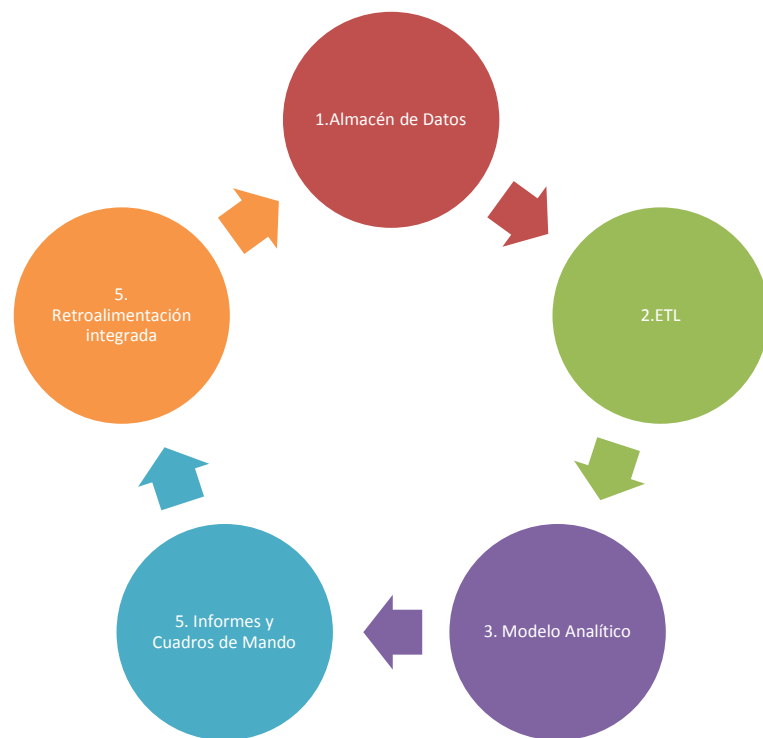


Figura 5: Secuencia del Proyecto

Para llevar a cabo la primera fase, nos encontramos con muchas metodologías de diseño y construcción de DW. Cada fabricante de software de inteligencia de negocios busca imponer una metodología con sus productos. Sin embargo, se imponen entre la mayoría dos metodologías, la de Ralph Kimball y la de Bill Inmon. Con el objeto de comprender la mayor diferencia entre estas dos metodologías, debemos explicar además de la noción de Data WareHouse y la idea de Data Mart [\[20\]](#). Un Data Mart (Kimball) es un repositorio de información, similar a un Data WareHouse, pero orientado a un área o departamento específico de la organización (por ejemplo Compras, Ventas, RRHH, etc.), a diferencia del Data WareHouse que cubre toda la organización, es decir la diferencia fundamental es su alcance.

Desde el punto de vista arquitectónico, la mayor diferencia entre los dos autores es el sentido de la construcción del DW, esto es comenzando por los Data Marts o ascendente (Bottom-up, Kimball) o comenzando con todo el DW desde el principio, o descendente (TopDown, Inmon). Por otra parte, la metodología de Inmon se basa en conceptos bien conocidos del diseño de bases de datos relacionales, la metodología para la construcción de un sistema de este tipo es la habitual para construir un sistema de información, utilizando las herramientas habituales, al contrario de la de Kimball, que se basa en un modelado dimensional (no normalizado)

Estas dos metodologías, marcarán el desarrollo del proyecto y dado que se implementará un sistema de BI orientado a ventas, es lógico pensar que se utilizará la metodología de Ralph Kimball y ciertamente se utilizará esta metodología para implementar un Data Mart de Ventas, pero al igual que en la elección del sistema, en la cual se ha optado por un sistema híbrido, la metodología de Bill Inmon debe orientar nuestro proyecto. Las empresas generan planes estratégicos, generalmente para los próximos cinco años, es labor de los responsables de cada departamento alinear sus propios planes estratégicos con los objetivos de la empresa. Esta percepción es la misma que se ha de tener en cuenta en el desarrollo del proyecto, si bien desarrollamos un Data Mart de Ventas, este ha de estar alineado con el Data Warehouse de la empresa, aunque este no esté aún desarrollado. La dirección de la empresa debe participar activamente en el desarrollo y validar que el Data Mart de ventas, responde a la estrategia corporativa.

2.2 Diseño

Cada una de estas fases necesita ser implementada sobre herramientas software concretas, con fin de adecuar la decisión de las herramientas necesarias para la implementación del proyecto, se hace necesario conocer los objetivos de cada una de estas fases.

Para comenzar correctamente el diseño los dos Roles tienen que trabajar de forma conjunta, para que principalmente el rol de TI, interiorice y comprenda el plan estratégico de la compañía, lo cual facilitará un diseño coherente durante las tres primeras fases.

2.2.1 Almacén de Datos

Se trata de sistemas pensados para almacenar grandes volúmenes de dicha información histórica, de manera que las operaciones de consulta que se realicen sobre ellos para dar respuesta a las preguntas estratégicas que nos podamos plantear, sean lo más eficientes posibles.

Una colección de datos orientados por tema, integrados, variables en el tiempo y no volátiles. Este se caracteriza principalmente por ser:

1. Integrado: los datos almacenados en el Data WareHouse deben integrarse en una estructura consistente, por lo que las inconsistencias existentes entre los diversos sistemas operacionales deben ser eliminadas. La información suele estructurarse también en distintos niveles de detalle para adecuarse a las distintas necesidades de los usuarios.
2. Temático: sólo los datos necesarios para el proceso de generación del conocimiento del negocio se integran desde el entorno operacional. Los datos se organizan por temas para facilitar su acceso y entendimiento por parte de los usuarios finales. Por ejemplo, todos los datos sobre clientes pueden ser consolidados en una única entidad (dimensión) del Data WareHouse. De esta forma, las peticiones de información sobre clientes serán más fáciles de responder dado que toda la información reside en el mismo lugar.
3. Histórico: el tiempo es parte implícita de la información contenida en un Data WareHouse. En los sistemas operacionales, los datos siempre

reflejan el estado de la actividad del negocio en el momento presente. Por el contrario, la información almacenada en el Data Warehouse sirve, entre otras cosas, para realizar análisis de tendencias. Por lo tanto, el Data Warehouse se carga con los distintos valores que toma una variable en el tiempo para permitir comparaciones.

4. No volátil: el almacén de información de un Data Warehouse existe para ser leído, y no modificado. La información es por tanto permanente, significando la actualización del Data Warehouse la incorporación de los últimos valores que tomaron las distintas variables contenidas en él sin ningún tipo de acción sobre lo que ya existía.

2.2.2 ETL

Las herramientas denominadas ETL (Extraction, Transformation and Loading), se utilizan para cargar el almacén de datos, necesitamos herramientas que nos permitan extraer los mismos desde distintos orígenes de datos, cuando se parte de bases de datos transaccionales, son heterogéneos, pero se hace necesaria la transformación, para poder cargar los datos de manera que responda al modelo analítico que se haya escogido.

1. Extracción: En nuestro caso práctico es necesario conectar solo una fuente, la conexión con el sistema transaccional. Pero de manera general en los procesos de extracción podemos encontrar todo tipo de fuentes.
2. Transformación: Debido a la naturaleza heterogénea de las fuentes y los datos podemos encontrarnos con problemas como:
 - a. Campos con longitudes diferentes en las distintas fuentes
 - b. Descripción inconsistente: Un mismo campo tiene distinto significado en las distintas fuentes
 - c. Distintas codificaciones para el mismo término, este problema no se presenta en nuestro caso práctico, debido a que solo existe una fuente, pero en cualquier caso si la fuente no está correctamente

normalizada, podemos encontrar el mismo problema a nivel de entidades.

- d. Valores nulos que suelen ser bastante frecuente, ya que no todos los valores extraídos tienen que ser obligatorios, por lo que por norma general se suele usar el valor -1 identificado como Indeterminado.
 - e. Adaptación de la reglas de integridad: Para que la integridad de los mismos se adapte a las reglas del almacén destino. Es aconsejable normalizar las relaciones con claves principales nuevas numéricas, estas permitirán un mejor rendimiento del modelo analítico, el join y group by son las operaciones que más afectan al rendimiento de una base de datos.
 - f. Formato: Por ejemplo el formato de una fecha puede no coincidir con el que pretendemos almacenar.
 - g. Seleccionar solo datos necesarios para el modelo analítico, se trata de un sistema para analizar, si no excedemos con la granularidad del sistema, el rendimiento se ve afectado y su propósito comprometido, ya si se permite al usuario final (usuario transicional), decidir al respecto, intentará resolver problemas funcionales en un sistema BI, por ejemplo Informe Intrastat de sistema ERP, ofrece un listado agrupado por país, cliente, factura, código intrastat del artículo de las ventas. Esta operación es puramente funcional, ofrecer al cliente sus ventas, agrupadas por factura y código intrastat.
3. Cargar los datos al almacén. Conectar con el repositorio del almacén de datos para cargar los datos. Antes puede ser necesario revisar la integridad de nuevo, calcular datos agregados, crear índices de las tablas destino,...

Es importante que las extracciones se realicen en el momento que la fuente soporte menos carga de usuarios, para garantizar el rendimiento del sistema transaccional del caso práctico.

Para aumentar el rendimiento en esta fase se puede dividir en 3 procesos claramente definidos:

1. Carga de Stage, es el único momento en el que se puede ver comprometido el rendimiento de la fuente de datos, a tal efecto las consultas de extracción tienen que ser rápidas, no tiene sentido sobrecargar una consulta sobre el sistema transaccional, cuando, podemos realizar a continuación las transformaciones necesarias para completar el modelo analítico
2. Carga de DWH, en esta fase se producirán las transformaciones necesarias para cargar los datos alineados con el diseño del modelo analítico, generación de claves principales nuevas, fusión de entidades de Stage, comprobación de integridad ...
3. Load: En esta última se procesara DWH dentro del modelo analítico

2.2.3 Modelos Analíticos (OLAP)

On-Line Analytical Processing, procesamiento analítico en línea, los cubos, las dimensiones y las jerarquías son la esencia de la navegación multidimensional de los modelos analíticos OLAP. Al describir y representar la información en esta forma, los usuarios pueden navegar intuitivamente en un conjunto complejo de datos.

Normalmente se usa un esquema de diseño en estrella, de copos de nieve y constelaciones de hechos [\[21\]](#). Este diseño es subyacente al software OLAP empleado, a continuación se puede observar en la Figura 6 un ejemplo de esquema de estrella.

MODELO ESTRELLA

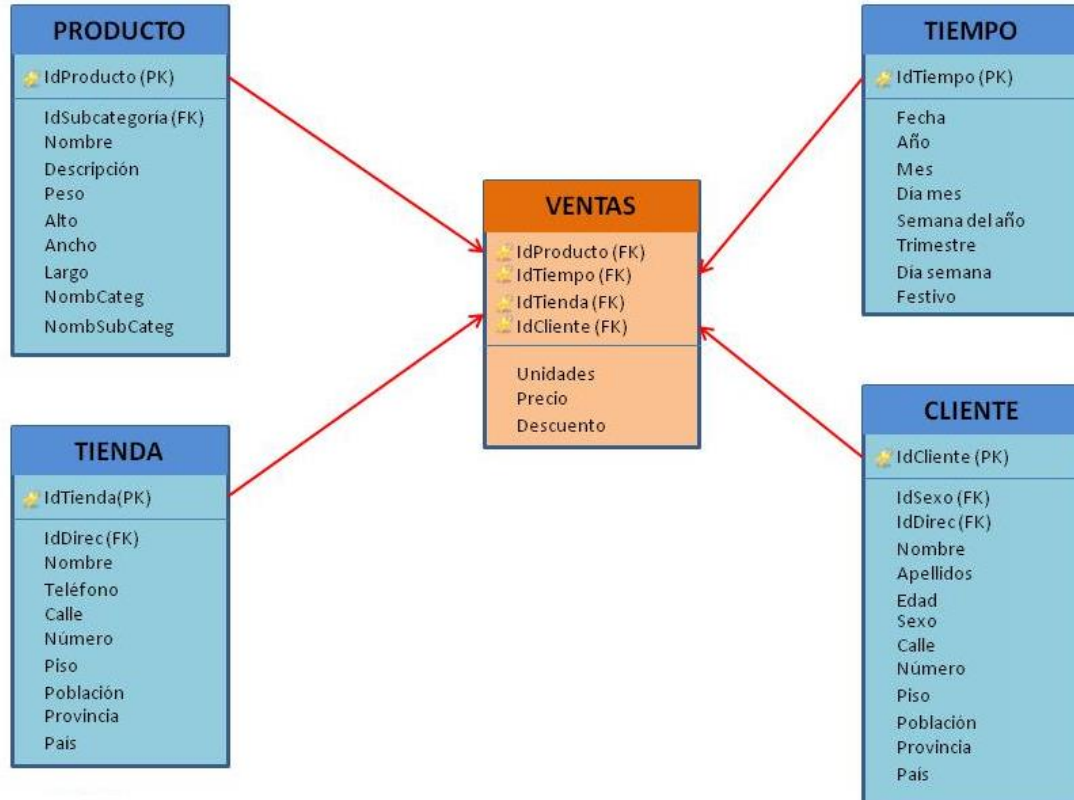


Figura 6: Esquema Estrella

En la Figura 8 podemos ver un ejemplo de esquema Copo de Nieve

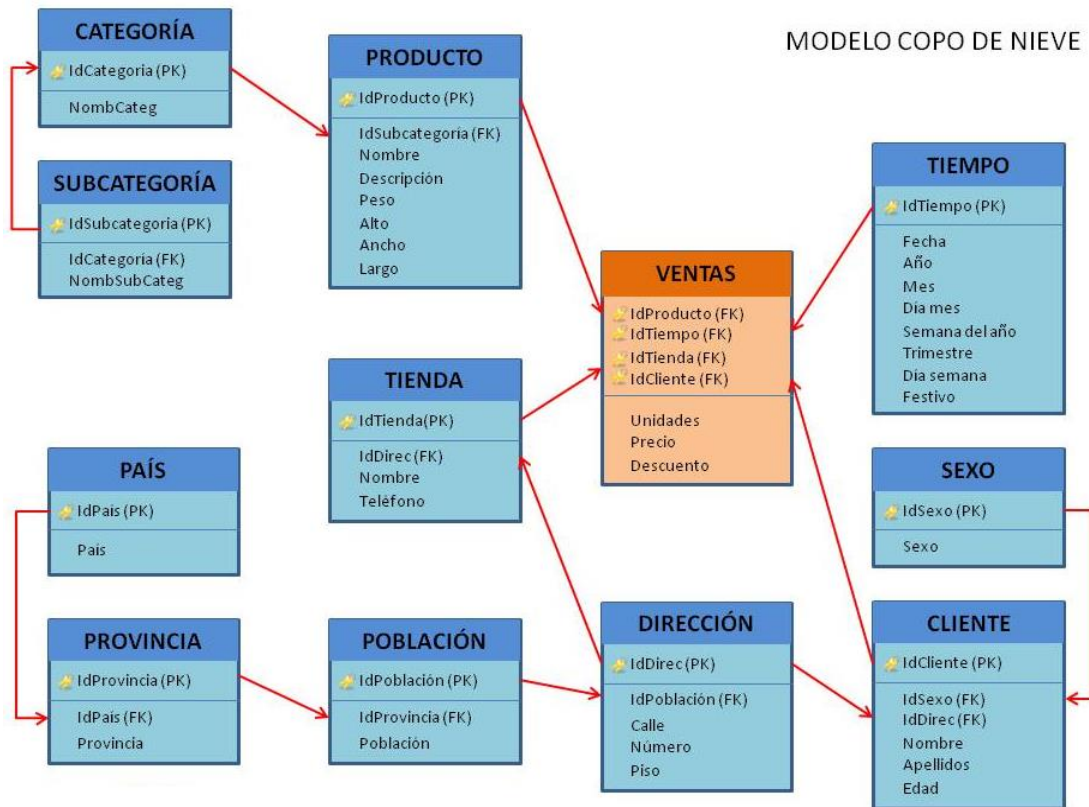


Figura 7: Modelo Copo de Nieve

Existen varias herramientas para realizar análisis de datos a través de un modelo de datos multidimensional, MOLAP (Multidimensional OnLine Analytical Processing, procesamiento analítico multidimensional en línea), ROLAP (Relational OnLine Analytical Processing, Procesamiento Analítico OnLine Relacional), HOLAP (Hybrid Online Analytical Process, procesamiento analítico en línea híbrido) [22], este último es un híbrido de los dos anteriores. La utilización de ROLAP o MOLAP es algo que hay que plantearse en la etapa del diseño físico. Esta decisión afecta a la manera en que la herramienta de explotación del Data Warehouse ataca a los datos. Al usar herramientas de Data Discovery es preferible utilizar ROLAP, ya que su compresión resulta más simple

de cara al usuario de negocio, permitiendo una rápida interiorización del esquema.

Como curiosidad, con Analysis services de SQL Server se puede elegir si el almacenamiento físico de los cubos se hace en ROLAP, MOLAP o HOLAP. Este es un caso especial, ya que este ROLAP se reduce a la implementación de los cubos sobre vistas indexadas de la base de datos en lugar de utilizar el sistema de ficheros. Tiene varias restricciones y el tiempo de procesamiento es mayor, aunque puede ahorrar espacio. Para mí este modo ROLAP de SSAS no es un ROLAP estándar, son sólo cubos OLAP implementados con vistas. Un ROLAP estándar se implementa sobre tablas físicas especialmente diseñadas siguiendo un modelo en estrella o un modelo en copo de nieve para obtener los mejores tiempos de respuesta y todas las ventajas del data Warehousing.

Esta es la última fase en la que el rol predomina, el usuario de TI completará el modelo de datos de la solución, a parte de los detalles técnicos de la fase, el usuario de TI tiene que tener muy en cuenta las necesidades analíticas del rol de Negocio a la hora de implementar los cubos.

2.2.4 Informes y Cuadros de Mando

En este punto el Rol de TI pierde el protagonismo y es el usuario de negocio es el que se responsabiliza de implementación de Informes y Cuadros de mando. Para llevar a cabo su tarea este ha tenido que participar activamente en la fase anterior, lo cual le facilita el conociendo necesario para el desempeño de su tarea.

Es importante comprender el término usuario de negocio, como un rol, de una o varias personas que toman decisiones estratégicas. Si bien tradicionalmente el usuario de negocio se identifica con un director general,

marketing o comercial (administración, compras, logística...), no hay que obviar que el comercial, el agente de marketing, el operador de compras... son a su modo un usuario de negocio, los cuales no han de excluirse del sistema de BI, simplemente se trata de usuarios con un nivel de acceso a datos menor.

Cada esquema estrella que implementamos en el sistema tiene imagen en una entidad del organigrama de empresa, el cual es muy difícil de cambiar físicamente, pero en un sistema de BI este puede ser remodelado en base a estrategias distintas, por ejemplo en la Figura 8 se puede ver un organigrama de empresa basado en departamentos.

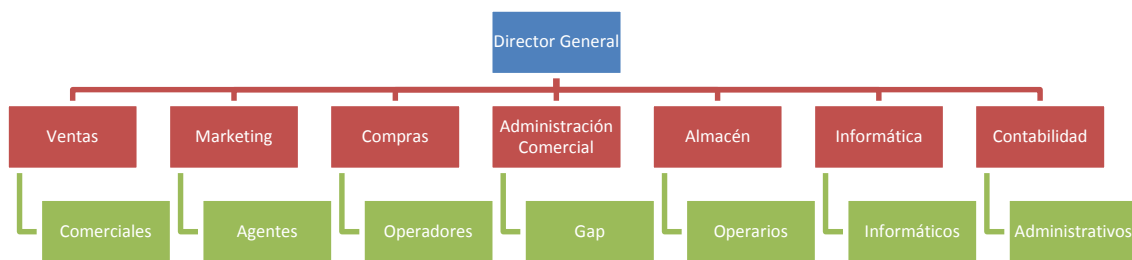


Figura 8: Organigrama por Ámbito

En el esquema generado para dar respuesta a las ventas dentro de nuestro sistema BI, podemos remodelar el organigrama de basado en departamentos a un organigrama basado en Roles o Funciones, obteniendo agrupaciones funcionales del esquema de ventas. En la Figura 9 no encontramos un ejemplo de remodelado funcional, donde los departamentos responsables directos de la venta de una empresa se reagrupan en el rol “fuerza de ventas”.

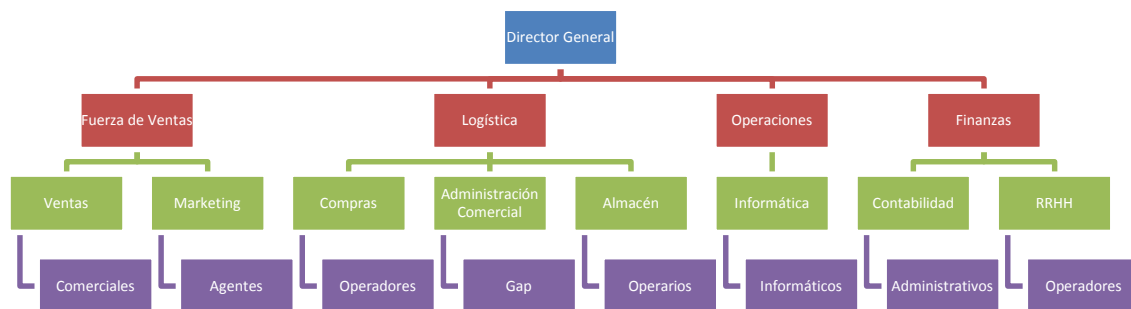


Figura 9: Organigrama Funcional

Realmente la información necesaria por Marketing es tan distinta de la necesaria por comercial, reagrupando áreas de la empresa encontramos nuevas realidades afines a su trabajo, basadas principalmente en sus relaciones.

Es muy importante que el usuario de negocio interiorice del modelo analítico diseñado en la fase anterior y la formación adecuada en la herramienta de Discovery Data seleccionada. Esta parte puede solapar con el punto anterior, para que una vez completo el modelo analítico, no se demore el sistema, a la espera de la formación.

El usuario de negocio genera a partir del modelo analítico informes funcionales que representan realidades estratégicas de la compañía de manera aislada, una vez se ha generado una respuesta individual a esta realidades, el usuario de negocio genera los Cuadros de Mando, lo cuales incorporan partes de los informes estratégicos más representativos.

Finalmente y a partir de esos cuadros el usuario de negocio, comprende claramente que indicadores son estratégicos en cuadro y genera los KPI Key Performance Indicator, conocido como indicador clave de desempeño o también indicador clave de rendimiento [\[23\]](#) que deben acompañar a los Cuadros de Mando.

En la Figura 10 se muestra la secuencia tradicional necesaria para la toma de decisiones.

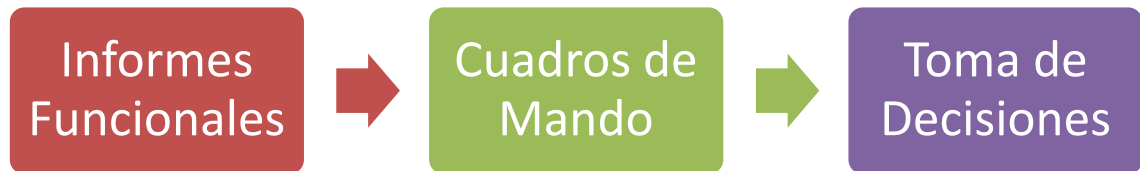


Figura 10: Secuencia de creación de Información

3. Retroalimentación integrada centrada en el usuario para la toma de decisiones

Completado el modelo híbrido, como punto clave del proyecto, se incorpora al descubrimiento de los datos, un elemento diferenciador, en el que la toma de decisiones, genera nuevo conocimiento en modelo datos. Hay que completar el proyecto añadiendo una nueva fase en la que el usuario final retroalimentará el sistema de BI.

Para la retro Alimentación del sistema, es necesario hacer uso de los informes y cuadros de mando generados, los cuales han de reanalizarse en base a la metodología DAFO, metodología empleada para la toma de decisiones de ventas. Este análisis consta de cuatro pasos principales:

1. Análisis Externo, también conocido como "Modelo de las cinco fuerzas de Porter", su objetivo será identificar a través de los informes y cuadros de mando generados en el sistema las distintas Amenazas y Oportunidades, al tratarse de un análisis sobre la venta, esta estará representada principalmente por tipos de clientes y la competencia.
2. Análisis Internos, su objetivo será identificar a través de los informes y cuadros de mando generados en el sistema, las Debilidades y Fortalezas principalmente de las familias de productos de la empresa.
3. De la combinación de los factores externos e Internos, surgen:
 - a. Las Potencialidades, obtenidas de la combinación de fortalezas con oportunidades, estas señalan las líneas de acción más prometedoras para la empresa.
 - b. Las Limitaciones, determinadas por una combinación de debilidades y amenazas, colocan una seria advertencia.
 - c. Los Riesgos (combinación de fortalezas y amenazas)
 - d. Los Desafíos (combinación de debilidades y oportunidades),

4. Para dar respuestas a estas combinaciones de factores, surgen distintas estrategias alineadas con el plan estratégico de la empresa. Es necesario la situación teniendo en cuenta la realidad particular de lo que se está analizando, las posibles alternativas (acciones) a elegir y las consecuencias futuras de cada elección (objetivos)

Una vez completado los cuatro pasos el usuario de negocio reconoce claramente las principales potencialidades, limitaciones, riesgos y desafíos, genera las estrategias necesarias, las acciones y objetivos esperado de dichas acciones. Todos estos elementos han de tener un marco temporal, particularmente las acciones han de tener una fecha de inicio y de fin, en cual si todo es correcto el objetivo debería haberse cumplido.

Con la información resultante, se debe retroalimentar el sistema ERP en forma de transacción, permitiendo a los usuarios funcionales participar de forma activa en la implementación de las acciones derivadas de la estrategia. Las cuales volverán nuevamente a través de un proceso ETL específico al sistema de BI.

Esta realimentación del sistema de BI, permite ver la evolución de las acciones frente a los datos reales, y finalizado el tiempo de implementación de la acción, el sistema permite que el usuario de Negocio:

- Verifica que la implementación de las acciones en el sistema transaccional.
- Contrastar los resultados esperados (los objetivos) con los obtenidos realmente.
- Generar Conocimiento, estas acciones resultantes de la estrategia, aportando un indicador o KPI, que indica en base a los objetivos marcados, el porcentaje de éxito de la acción. Este indicador facilita a toda la organización una experiencia contrastada, que genera una

base de conocimiento estratégico sobre el proceso de Toma de Decisiones.

Esta nueva fase convierte el modelo secuencial actual del Data Discovery que se puede observar en la Figura 11

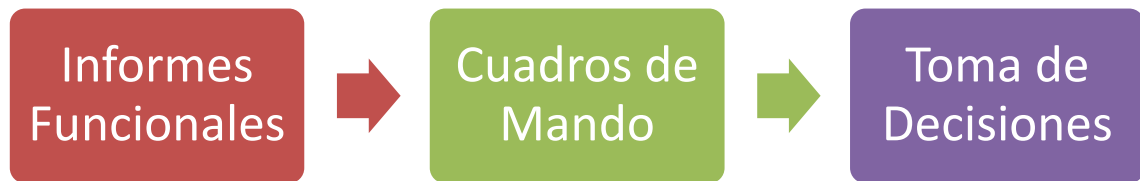


Figura 11: Modelo Secuencial

El modelo secuencial se convierte en un modelo cíclico, que crea conocimiento nuevo, en base a las decisiones adoptadas, aportando al usuario de negocio la base de conocimiento obtenida de decisiones anteriores. La secuencia resultante se puede observar en la Figura 12.

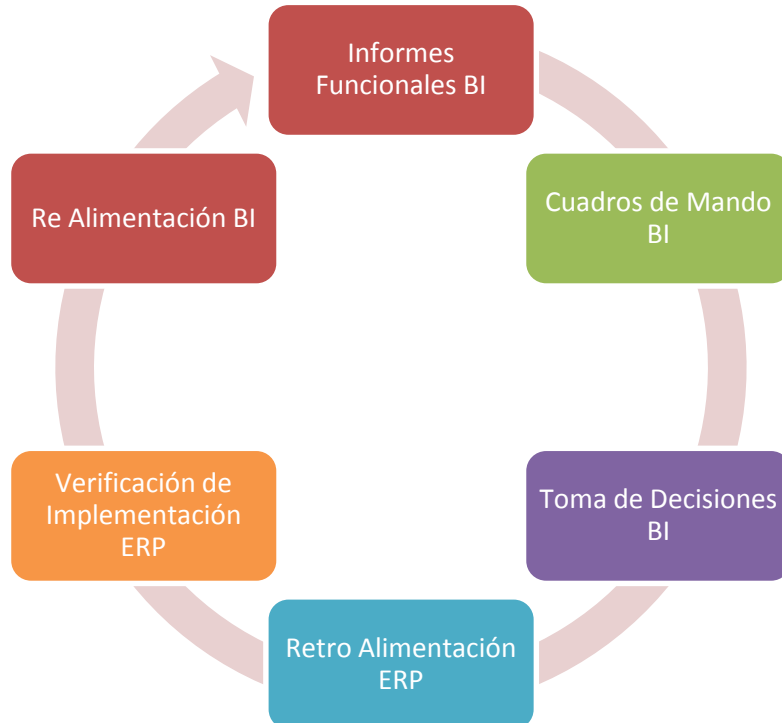


Figura 12: Modelo Cíclico

La toma de decisiones es un proceso cognitivo [24] que requiere que el usuario de negocio procese la información a partir de la percepción, el conocimiento adquirido (la experiencia) y características subjetivas que permiten valorar la información. Para la consecución de este diseño se requiere habilidades como el aprendizaje, razonamiento, atención, memoria, resolución de problemas, toma de decisiones y procesamiento del lenguaje. Todos estos elementos han de tener su imagen en el sistema de toma de decisiones implementado.

En el sistema secuencial, la información se almacena en la memoria del usuario, mientras que en el sistema propuesto, dispone el conocimiento de la toma de decisiones, almacenada sobre un soporte físico (una base de datos). Como principales consecuencias de este cambio de soporte de almacenamiento, donde el usuario es el soporte, se eliminan los siguientes riesgos inherentes:

- a. La memoria no garantiza la recuperación de la información en el momento preciso
- b. La experiencia colectiva (la toma de decisiones de otros usuarios) no se registra en la memoria individual, y en consecuencia no puede usarse directamente por el individuo, que depende de que esta sea compartida correctamente.
- c. Continuidad de negocio, cualquier usuario de negocio puede ceder en su actividad por múltiples motivos, y su experiencia desaparece con él.

Con el planteamiento Cíclico del proceso de toma de decisiones, se pretende evolucionar a partir de sistemas de inteligencia de negocio a un sistema que bien podría denominarse sistema cognitivo de negocio.

El descubrimiento de la información propuesto en modelo de Data Discovery, donde el usuario negocio:

1. Descubre los hechos relevantes en base a informes y cuadros de mando
2. Toma las decisiones en base al descubrimiento

Pasamos a un modelo donde el usuario de negocio:

1. Descubre los hechos relevantes en base a informes y cuadros de mando
2. Toma las decisiones en base al descubrimiento
 - a. Clasifica los hechos afectados en base a Análisis DAFO
 - b. Diseña estrategias en base al análisis y el éxito de estrategias empresariales anteriores
 - c. Se elabora acciones limitadas en el tiempo que respalden la estrategia diseñada.
3. Retro Alimentación el sistema de Información (Toma de decisiones→ERP→BI).
4. Verifica el éxito o fracaso de su estrategia

4. Herramientas

Para la implementación del proyecto, se han tenido en cuenta herramientas del mercado que permitan la generación del modelo propuesto, en el que se hace necesario:

1. Un almacén de datos
2. Procesos ETL (Un fuente de datos transaccional/ERP)
3. Un modelo OLAP
4. Herramientas de procesamiento de modelos OLAP para la generación de informes y cuadros de mando.

Para implantar dicho modelo se ha optado por dividir el proyecto en dos bloques.

1. Almacén de Datos, ETL y sistema OLAP
2. Herramientas de procesamiento de modelos OLAP

4.1 Almacén de Datos, ETL y sistema OLAP

En el mercado actual podemos encontrar suites que contengan todos estos elementos, facilitando la implementación de un sistema de inteligencia de negocio homogéneo. Estas herramientas que permiten la generación de un proyecto completo, abordado de manera integral.

4.1.1 Cuadrante Mágico de Gartner para plataformas de BI y Analíticas 2015

En la Figura 13 se muestra el cuadrante mágico de Gartner [\[25\]](#) para plataformas de BI y analíticas a Febrero 2015.



Figura 13: Informe Gartner 2015 plataformas de BI

Como se puede observar continúan siendo líderes las plataformas tradicionales:

Microsoft, SAS, SAP, IBM, Oracle, MicroStrategy, Information Builders

Las herramientas de BI: Qlik y Tableau permanecen en el cuadrante de líderes. Es importante recordar que estas herramientas no ofrecen soluciones de

Data Warehousing y permiten conectarse directamente a un Data Warehouse o cualquier otra fuente de datos y demandan de bastante memoria RAM en los servidores. Sin embargo, su costo de licenciamiento comparado con las plataformas tradicionales es bajo.

Sobre las herramientas Open Source: Pentaho continúa en el cuadrante de Jugadores de Nicho (Niche Players) y se anunció éste año la adquisición por parte de Hitachi Data Systems. Jaspersoft fue adquirido por Tibco Software que se encuentra en el cuadrante de Visionarios.

Referencias.

Explicación del Cuadrante [\[26\]](#)

Informe Oficial de Gartner [\[27\]](#)

4.1.2 Herramienta seleccionada

Para la implementación del proyecto se ha tenido en cuenta el origen de datos, al ser una fuente Microsoft (Microsoft Dynamics AX 2009 sp1 sobre Microsoft SQL 2008).

El entorno Microsoft ofrece un motor de base datos que actúa como almacén de datos (Microsoft SQL 2014), una es una plataforma para la integración de datos y sus de flujos de trabajo (SQL Server Integration Services, SSIS) y herramientas OLAP (Microsoft SQL Server Analysis Services). [\[28\]](#)

La suite de Microsoft contiene todos los elementos necesarios para la implementación de un proyecto de BI, a nivel de reporting se puede utilizar SQL Server Reporting Services, pero dado que el modelo propuesto sustituye esta fase por herramientas de Data Discovery, no será necesario en el proyecto.

El sistema de Microsoft ofrece los servicios necesarios para nuestro proyecto, para configurar dichos servicios pone a disposición del usuario de TI, su framework más característico, Visual Studio 2013, el cual incorpora Microsoft SQL Server Data Tools - Business Intelligence [\[29\]](#) (en febrero del 2014 esta herramienta quedo liberalizada para la implementación de proyectos de BI de forma Gratuita). Estas herramientas junto con el Visual Studio nos permiten implementar las tres fases del proyecto en un único entorno.

4.2 Herramientas de procesamiento de modelos OLAP

La fase de generación de informes y cuadros de mando, será implementada con herramientas Data Discovery, si bien el mercado existen distintas opciones, se ha seleccionado Excel 2013 entre todas ellas, por los siguientes motivos:

1. Integración, partiendo de un ERP Microsoft, Un sistema de BI Microsoft, Un sistema de Data Discovery Microsoft debe garantizar la compatibilidad de todos estos sistemas de manera nativa.
2. Coste de Aprendizaje, el usuario de negocio actual suele conocer Excel, no es necesario invertir en costosas formaciones para su uso.
3. Coste de la Herramienta, el paquete office es el líder indiscutible del mercado, y un gran número de organizaciones ya disponen del software.

La información desde sus orígenes se basa en dos roles principalmente consumidor y productores. Las necesidades de lectura, por norma general en sistemas de análisis, mucho más altas que las de escritura. Estas necesidades evolucionan por factores normalmente externos (mercado, competencia, clientes...), la información crece de manera muchas veces exponencial y estos cambios se producen normalmente mucho más rápido que las capacidades evolutivas de los propios sistemas. El BI como tal tampoco tiene capacidad evolutiva, pero al contrario que los informes tradicionales, puede evolucionar mucho más rápido.

Reporting → BI Tradicional → Auto Discovery

¿Qué ofrece Auto Discovery al BI?

Principalmente fluidez, la forma en que actualmente se genera la explotación y visualización en BI tradicional, impide que el usuario realice preguntas fuera de los modelos de análisis de la empresa de manera fluida.

La información se genera correctamente, permite análisis fuera de los modelos (cuadros de mando estándar), permitiendo detectar situaciones no contempladas en dichos modelos., interactuando con herramientas de data mining y planteado posibles escenarios hipotéticos que ayudarán a la toma de decisiones.

La importancia del Auto Discovery reside principalmente en:

1. La información es gestionada por quien la explota
 - a. La información se obtiene más rápido: debido principalmente a que la necesidades puntuales de información pueden ser resueltas por los propios usuarios de negocio
 - b. Autodescubrimiento, las preguntas las realizan los propios usuarios de negocio y las resuelven.
 - c. Ante necesidades altas de información, no se generan cuellos de botella en los departamentos de TI.
2. Separar BI en dos capas: Construcción (TI) y Explotación (Usuarios de Negocio), delimitando claramente las responsabilidades. TI da servicio al sistema no al usuario, el sistema da servicio al usuario.

El autoservicio supone un cambio en la cultura de la empresa, donde tradicionalmente el usuario requería el reporting y el BI tradicional de los equipos

de TI. Con esta “filosofía” el usuario deja de ser meramente un consumidor de información, pudiendo ejercer funciones de constructor, en entornos controlados.

4.2.1 Data Discovery

El campo relativamente nuevo del descubrimiento de datos está creciendo a un ritmo sorprendente, tanto en el número de empresas competidoras como en la cuota de mercado de estas empresas focalizadas en el análisis de datos es cada vez más grande. Según 2011 Descubrimiento de Datos Herramientas de Gartner informa "el descubrimiento de datos será un mercado de US \$ 1 mil millones".

Sin embargo, es un mercado relativamente pequeño en comparación con las tecnologías más antiguas, como la inteligencia de negocio (el número de herramientas de inteligencia de negocios supera con creces el número de herramientas de descubrimiento de datos).

Las herramientas de descubrimiento de datos en general, trabajan junto a sus usuarios para descubrir conocimientos ocultos en los datos de una forma interactiva y visual. Debido a esto, casi se podría considerar herramientas de descubrimiento de datos para ser herramientas de inteligencia de datos. Herramientas de descubrimiento de datos tienden a ser más visuales e interactivas que los informes tradicionales. Emplean nuevos métodos radicales de visualización de datos - tablas, gráficos, infografías, etc.

Estas herramientas permiten mostrar los resultados y descubrir al usuario a nuevos conocimientos e ideas. De hecho, la búsqueda de datos a menudo se ha referido como "data mining visual".

4.2.1.1 Los contendientes

Todavía no hay mucha competencia en este campo, pero startups se anuncian de forma casi semanal. Muchos de los vendedores les encantarían

estar asociados con el descubrimiento de datos, pero son realmente los proveedores de inteligencia de negocios disfrazados.

4.2.1.2 *Experiencia del usuario*

Aquí es donde los niños son separados de los hombres. Las propuestas de valor reales para herramientas de descubrimiento de datos se encuentran en su experiencia de usuario y la esperanza que ofrecen para hacer la búsqueda de datos a disposición de un público más amplio que la inteligencia de negocio tradicional.

Una de las cosas clave que los usuarios quieren de herramientas de descubrimiento de datos es la capacidad para generar efectos visuales significativas e interesantes de sus datos. Esto es, naturalmente, donde la mayoría de productos brillan, en la Figura 14 se puede ver el cuadrante de Visualización según el Informe Gartner del 2014.

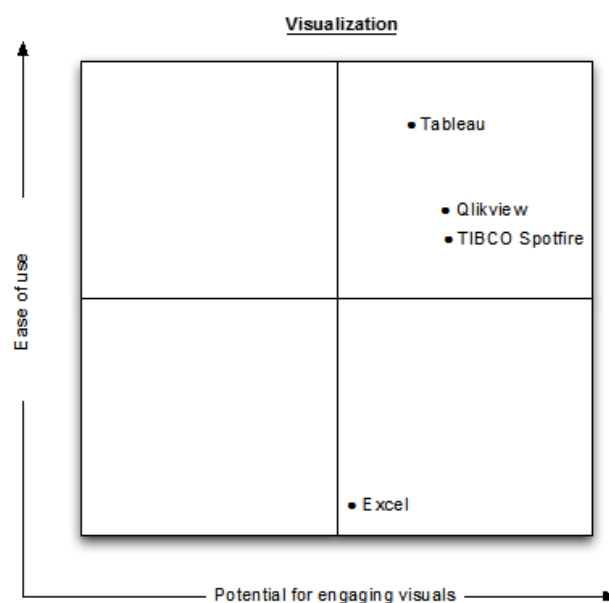


Figura 14: Informe Gartner 2014 Data Discovery Visualización

4.2.1.3 Movilidad

La experiencia móvil es cada vez más importante como los dispositivos móviles crecen en popularidad y los usuarios de descubrimiento de datos los utilizan en lugar de las estaciones de trabajo tradicionales. Si una pregunta es lo suficientemente importante, puede merecer sentarse en su computadora para hacer algunos análisis de investigación con fuentes y definiciones de datos, pero lo que realmente queremos es ser capaz de obtener nuestras respuestas donde y cuando los necesitamos. Y hoy eso significa en nuestros dispositivos móviles, ya sea nuestras tabletas o teléfonos inteligentes, en la Figura 15 se puede ver el cuadrante de Movilidad según el Informe Gartner del 2014.

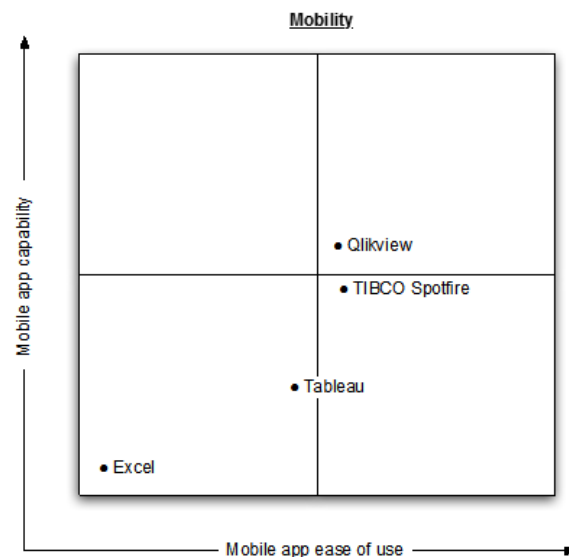


Figura 15: Informe Gartner 2014 Data Discovery Movilidad

4.2.1.4 Perfil de los Usuarios

Es evidente que la experiencia del usuario y la facilidad de uso tienen una gran influencia en que el usuario. Un ejecutivo no puede esperar para aprender

a instalar y utilizar el plugin de PowerPivot para Excel, por ejemplo. El usuario de destino para cada una de estas ofertas puede variar mucho. Esto debería ser una consideración importante cuando se decide qué solución es la mejor opción para sus necesidades, en la Figura 16 se puede ver el cuadrante de Facilidad de Uso por parte del usuario según el Informe Gartner del 2014.

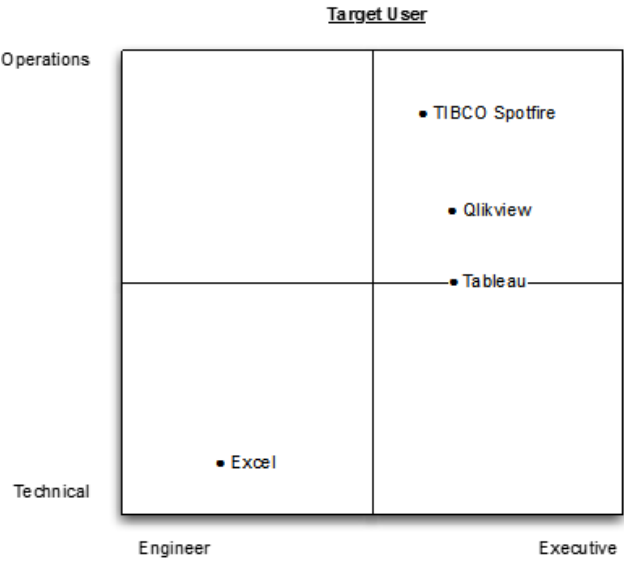


Figura 16: Informe Gartner 2014 Data Discovery Facilidad de uso por parte del usuario

Requisitos de Conocimientos técnicos, en la Figura 17 se puede observar, la necesidad de Conocimiento Técnico según el informe Gartner del 2014.

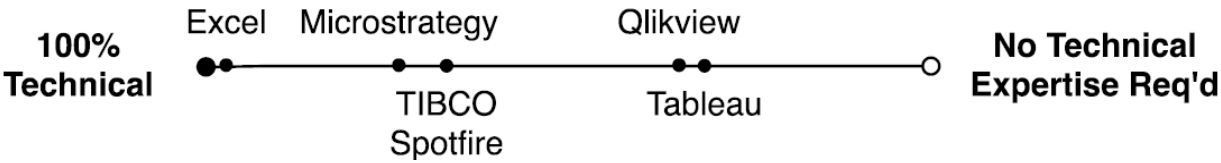


Figura 17: Informe Gartner 2014 Data Discovery conocimiento técnico

4.2.1.5 *Analítica del músculo*

A falta de un término mejor, la analítica del músculo es la destreza analítica de los productos. Algunas de las herramientas ofrecen una visión analítica muy profunda, proporcionando capacidades de predicción y análisis estadísticos, mientras que otros sólo hacen un buen trabajo de mostrar sus datos. Esto puede o no puede ser muy importante, en función del tipo de inteligencia de datos que está analizando, en la Figura 18 se puede ver el cuadrante de Escalabilidad según el Informe Gartner del 2014.

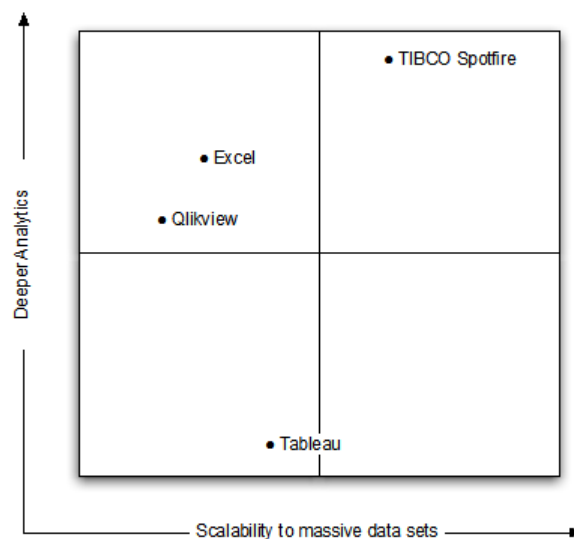


Figura 18: Informe Gartner 2014 Data Discovery Escalabilidad

En un intento por capturar y cuantificar el espíritu de descubrimiento visual de enganche de datos, potenciales para el descubrimiento de la visión y la facilidad de uso, se ha recopilado los factores en este informe para el siguiente resumen, el cuadrante de Liderazgo. Este cuadrante mide las dos características más importantes de una herramienta de descubrimiento de datos:

1. La calidad de la visión que se puede obtener a partir de ella.
2. La facilidad de obtener ese conocimiento.

En la Figura 19 se puede ver el cuadrante de Capacidad para entender las cosas con claridad y rapidez según el Informe Gartner del 2014.

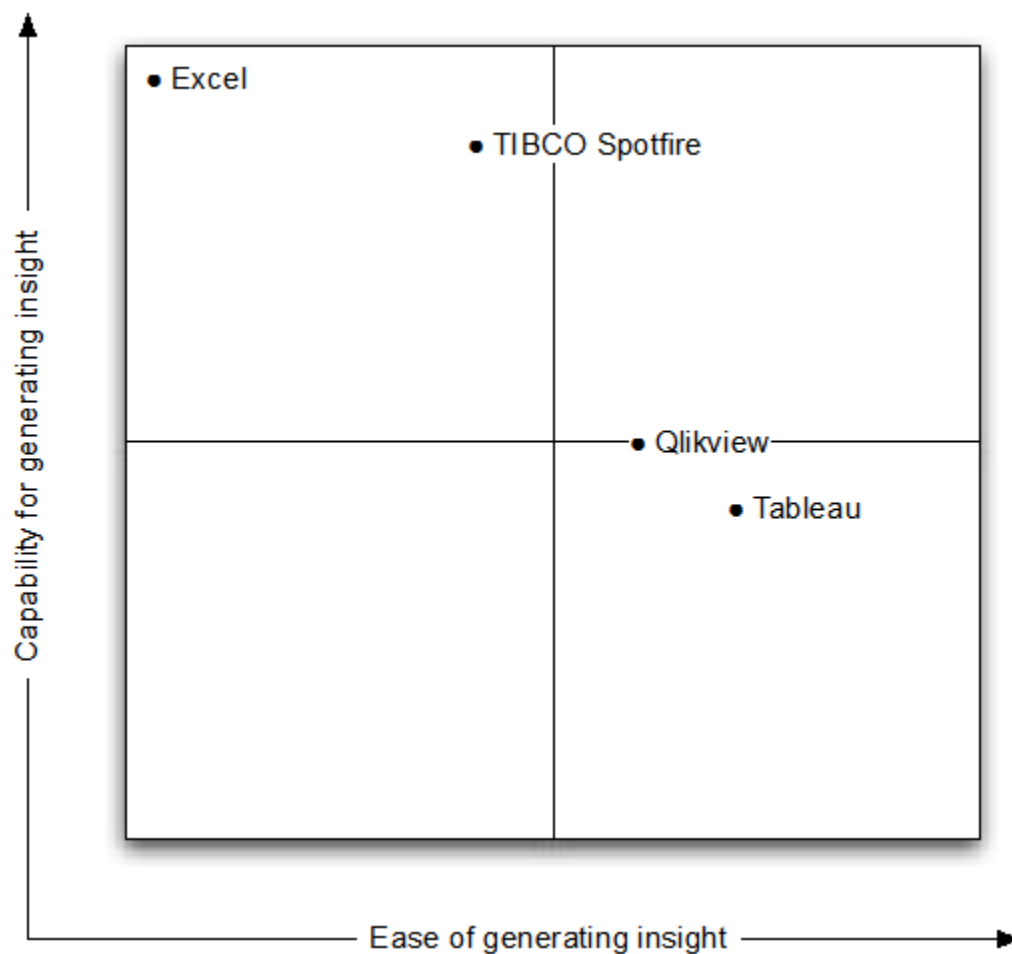


Figura 19: Informe Gartner 2014 Data Discovery generación de ideas

* Los datos fueron obtenidos de <http://www.applieddatalabs.com> este sitio ha desaparecido por lo que no se ha incluido en las referencias del proyecto

4.2.2 Herramienta seleccionada

Para la implementación del proyecto se ha tenido en cuenta principalmente la formación del usuario de negocio, el cual en un entorno Microsoft, aporta un conocimiento elevado del paquete office, en concreto el programa Excel 2013. También se ha tenido en cuenta que el coste de licenciamiento es 0, ya que la empresa dispone del paquete office.

Microsoft Excel es una herramienta difícil de manejar, pero con gran capacidad para el análisis. Dado que la empresa tiene experiencia previa en su uso, esta dificultad queda compensada.

Los cuadros CMI y KPI que permite Excel, son muy limitados en comparación con sus competidores directos, en caso de necesidad Microsoft aporta una herramienta superior para la generación de cuadros, Power BI, la cual es un complemento integrado en Excel 2013, que se puede adquirir bajo otro tipo de licenciamiento. Pero dado que el objetivo último del proyecto no son los Cuadros de Mando, Excel 2013 cubrirá sobradamente las necesidades asociadas al proyecto.

5. Caso Práctico

El marco del proyecto se centrará en las operaciones de venta de una empresa, a tal efecto se creará un modelo híbrido de BI, mezclando el BI más tradicional con Data Discovery. Para las tres primeras fases se ha utilizado el BI tradicional, el cual aporta la seguridad y rendimiento requeridos para un proyecto empresarial, las cuales serán extraídas a través de procesos ETL a un Data Mart sobre una base de datos implementada sobre un esquema en estrella. El Data Mart servirá el modelo analítico de las ventas que permitirá el DATA DISCOVERY. A partir de los análisis obtenidos por los usuarios se definirán los KPI e informes necesarios para la toma de decisiones, estos KPI e Informes estarán centrados en el análisis de los valores extremos de las muestras, permitiendo identificar y clasificar las debilidades y fortalezas de la venta en sí.

Las tres primeras fases han sido implementadas a través de Microsoft SQL Server Data Tools - Business Intelligence para Visual Studio 2013, el cual ofrece un entorno que permite todo el desarrollo del proyecto, donde se puede implementar el Data Mark directamente a través del Server Explorer (Figura 20)

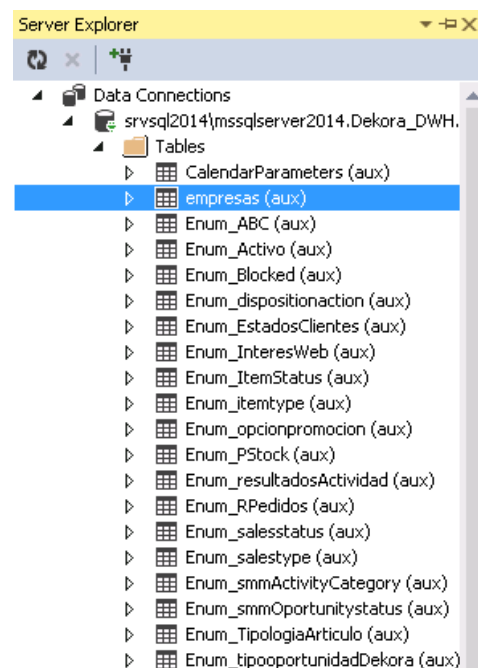


Figura 20: Server Explorer Visual Studio 2013

ETL a través de SSIS Packages, en la Figura 20 se puede observar el ETL de ventas implementado sobre Visual Studio 2013 con Data Tools.

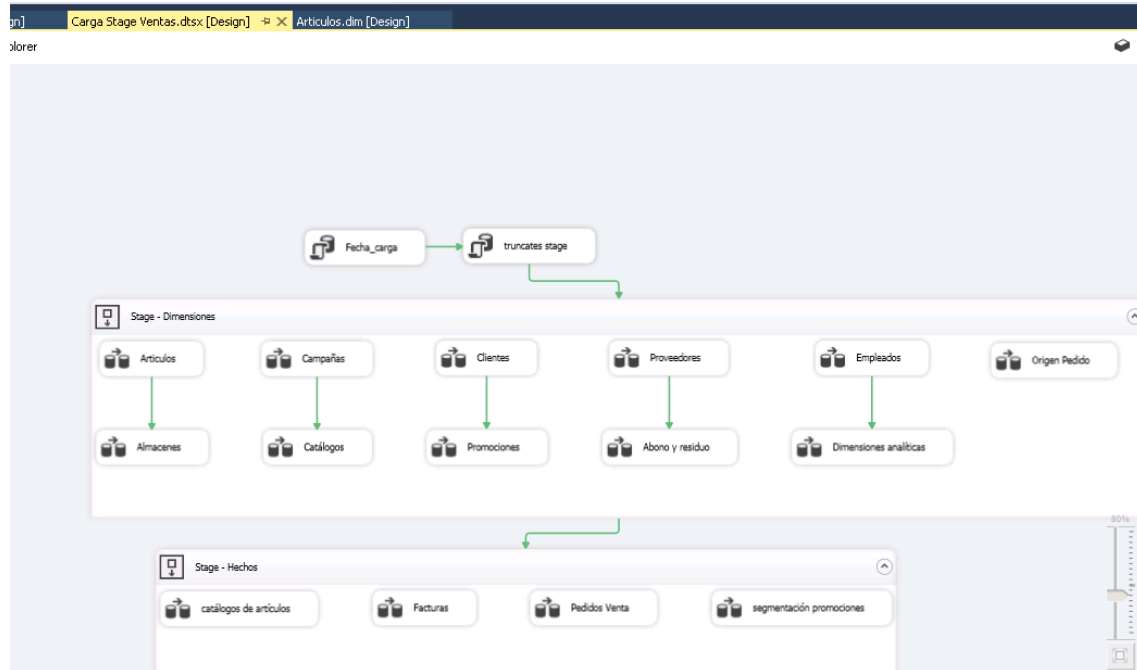


Figura 21: SSIS/ETL ventas (Extracción a Stage)

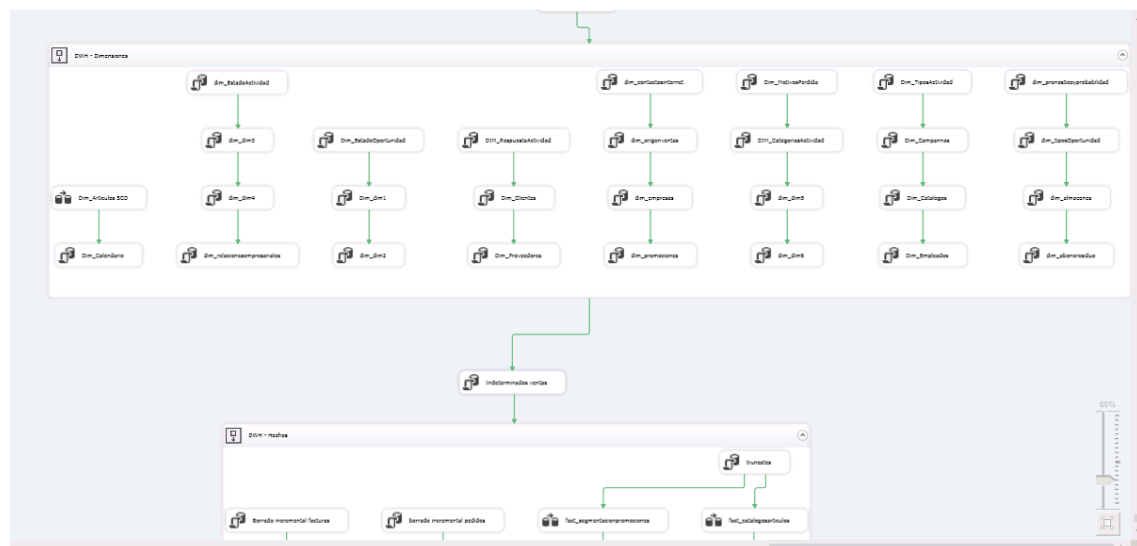


Figura 22: SSIS/ETL ventas (Transformación y Carga DataWareHouse)

Gestión de seguridad a nivel de dimensiones o registros, integrados con el directorio activo [30], los roles y la seguridad se implementan en la solución de manera nativa, tal y como se observa en la Figura 23.

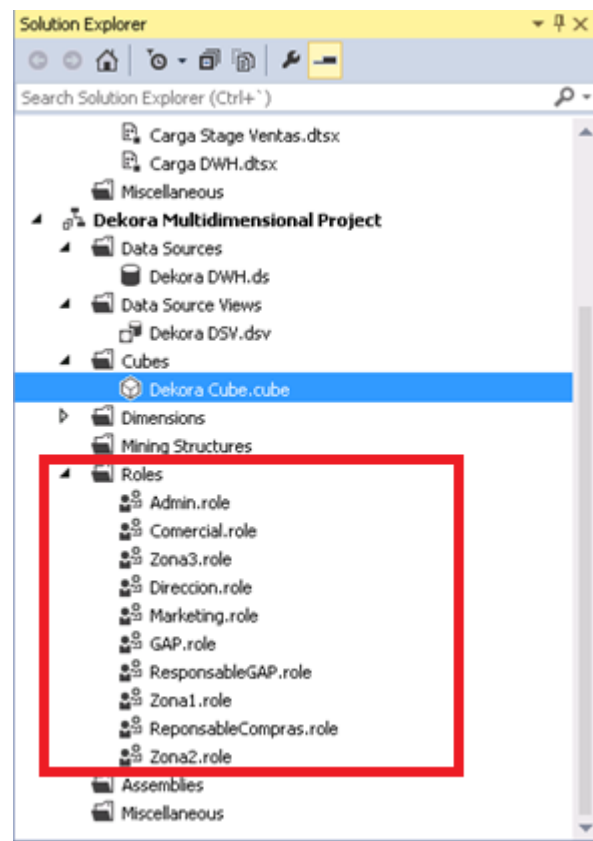


Figura 23: Solution Explorer (Roles de Seguridad)

Diseño OLAP, El paquete Data Tools permite la generación de los cubos en el mismo entorno, en la Figura 24 se puede observar el interface gráfico que facilita Data Tools, para este fin.

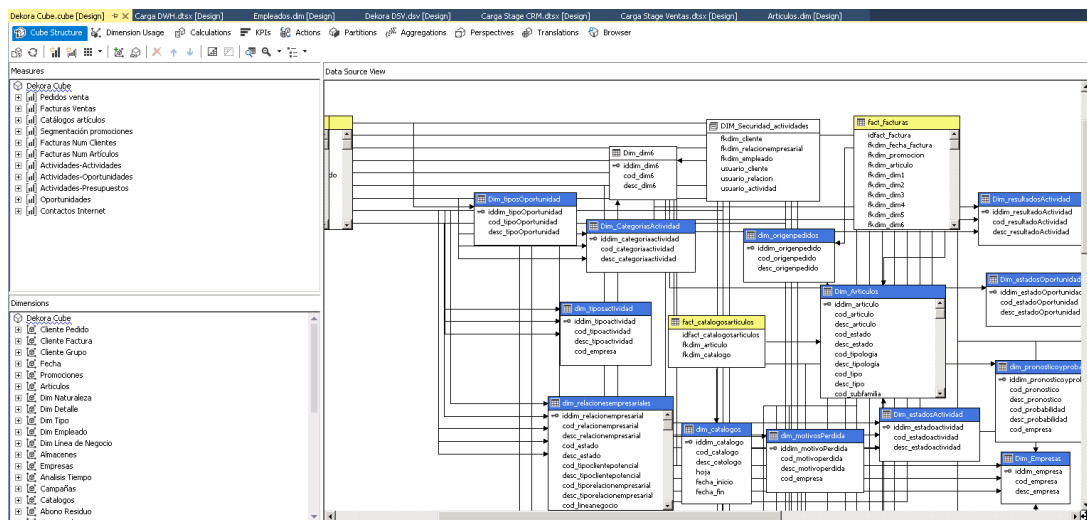


Figura 24: Modelo OLAP

Permitiendo otras opciones relacionadas con el modelo OLAP

Que dimensiones puede ver cada rol, medidas calculadas, KPI's, Acciones, Partition ... , las posibilidades de la herramienta Data Tools, tienen en cuenta casi todos los aspectos necesarios para la implementación de un Data Warehouse completo, en la Figura 25 se puede ver un ejemplo de otras operaciones que facilita el paquete Data Tools.



Figura 25: Opciones diseño OLAP

El paquete Microsoft SQL Server Data Tools - Business Intelligence para Visual Studio 2013 permite gestionar el proyecto en todos los ámbitos del BI e incluso Data Mining [\[31\]](#) o minería de datos integrado

A continuación detallamos la implementación de las distintas partes del caso práctico. Comenzaremos detallando el Almacén de datos para después ir completando las distintas fases del proyecto.

5.1 Almacén de Datos

Los almacenes de datos tienen su origen en el BI tradicional, y permiten separar la explotación de las bases de datos funcionales (ERP) de las bases de datos de análisis (BI), usar un almacén de datos corporativo facilita la gestión de la seguridad y el rendimiento.

Para un correcto análisis de las ventas hay que identificar claramente las unidades de medidas y las dimensiones necesarias para tal efecto, las estructuras de venta en ERP actual suelen ser muy similares, y tienen como eje central las facturas, albaranes y pedidos, estas tablas serán la base de las tablas de hecho del Data Warehouse. A nivel de dimensiones resulta evidente la utilización de artículos y clientes, pero dado que el problema no puede reducirse a estas dos entidades, se han incluido otras dimensiones de interés, proveedores (todo artículo tiene proveedor), Dimensiones analíticas (usadas en contabilidad para clasificar las entidades maestras), Promociones (los artículos son promocionados desde el departamento de marketing), Almacenes (La venta se produce desde un almacén concreto), Catálogos (los artículos pertenecen a catálogos) y Campañas (ciertos artículos responden a necesidades temporales del negocio). Toda esta información es obtenida en base a tres esquemas fundamentales:

5.1.1 Esquema Aux

En este esquema se guardan datos necesarios para las diferentes transformaciones y que no se recargan frecuentemente así como tablas administrativas del Data Warehouse:

- Parámetros

Campo	Tipo de datos	Descripción
-------	---------------	-------------

FechaInicioDWH	Date	Primera fecha para la que se cargarán datos en el almacén de datos.
FechaUltimaCarga	DateTime	Fecha en la que se lanzó el ETL de actualización por última vez.
CargaCompleta	Integer	Si este campo tiene el valor 1 se realiza una carga completa de transacciones en el DWH desde la Fecha de Inicio indicada en el campo anterior, borrando previamente todos los registros existentes. Si este campo tiene el valor 0, la carga de transacciones es incremental, en función del tipo de recarga indicado para cada tabla de hechos.

5.1.2 Esquema Stage

Este esquema recibe los datos desde la base de datos operacional. En él se almacena la precarga de datos y desde él se realizan las transformaciones necesarias en los datos. Este esquema realiza la extracción de los datos de manera eficiente, afectando lo mínimo al sistema de explotación ya que la extracción no contiene ninguna transformación y se realiza de manera selectiva para evitar más de dos join (unión de tablas) anidados, esta práctica es imprescindible de cara a obtener un alto rendimiento en la extracción de datos, ya que generalmente en un sistema de explotación ERP, las relaciones cumplen al menos la 5 forma normal, y la extracción directa puede ser impracticable. En

algunos casos como las relaciones entre entidades hacen uso de claves ajenas muy costosas en términos de rendimiento. Ejemplo facturas y líneas, que por norma general se relacionan a través del identificador de factura y la fecha de factura, un String y un Date.

Las tablas cargadas en el Stage tendrán un periodo de recarga dependiendo de la información que contengan:

- **Maestros.** Se recargan completos sin acotar fechas
- **Transacciones.** Se recargan desde una fecha especificada por el usuario.

A continuación se describen las tablas del Stage y cuál es su correspondiente origen en AX.

Sólo se traerán al stage los campos necesarios para el DWH.

Tabla Stage	Tabla AX correspondiente	Descripción tabla
Custtable	CUSTTABLE	Maestro de clientes
Inventtable	INVENTTABLE	Maestro de artículos
Vendtable	VENDTABLE	Maestro de proveedores
Inventlocation	INVENTLOCATION	Maestro de almacenes
Dimensions	DIMENSIONS	Maestro de dimensiones analíticas
Campaign	CAMPAIGN	Maestro de campañas
Promotion	PROMOTION	Maestro de promociones
Catalog	CATALOG	Maestro de catálogos
CatalogDate	CATALOGDATE	Asociación de campañas a intervalos de fechas anuales.
ItemCatalog	ITEMCATALOG	Asociación de catálogos a artículos

Tabla Stage	Tabla AX correspondiente	Descripción tabla
CampaignCust	CAMPAIGNCUST	Segmentación de promociones por clientes
CustInvoiceJour CustInvoiceTrans	CUSTINVOICEJOUR CUSTINVOICETRANS	Cabeceras y líneas de facturas de ventas
SalesTable SalesLine	SALESTABLE SALESLINE	Cabeceras y líneas de pedidos de ventas
Reason	REASON	Tipificaciones de situaciones de análisis
Solution	SOLUTION	Tipificaciones de las soluciones información de análisis

5.1.3 Esquema DWH

Una vez completado el esquema Stage, se transformaran los datos en al esquema DWH que contiene la información consolidada que procesará el sistema OLAP. Según su nomenclatura podemos distinguir dos tipos de tablas:

- **Tablas de dimensiones.** Son las tablas que almacenan los datos de los diferentes ejes de análisis, vienen precedidas por el prefijo “Dim_”.
- **Tablas de hechos.** Almacenan la información en su correspondiente nivel de detalle, vienen precedidas por el prefijo “Fact_”.

La actualización se realizará de diferente forma para dimensiones y hechos:

- **Tablas de dimensiones.** La actualización se realiza comparando qué registros son nuevos y en caso de existir el registro se actualizan sus características.
- **Tablas de hechos.** En este caso no hay comparación, se borra de las tablas de hechos el periodo a actualizar y se recarga.

Uno de los pasos más importantes de la carga es la conversión de claves operacionales en claves internas del DataWareHouse. Cualquier clave primaria en los sistemas operacionales tendrá una nueva clave en el DataWareHouse, convirtiéndose así en una clave alternativa.

Esta conversión se realiza cruzando los datos de las dimensiones con las tablas de Staging, por lo que se creará un miembro “Indeterminado” en cada una de las dimensiones para evitar la pérdida de datos. Cualquier clave no reflejada en un maestro se catalogara de esta manera.

Esta conversión supone un cambio en las uniones de las tablas, ya que a partir de este momento, estas se relacionaran a través de un identificador BIG INT, lo cual afectará muy positivamente al rendimiento.

A continuación se detallan las tablas del esquema DWH y su correspondencia con el Stage, en este esquema hay dos dimensiones que no tienen su origen directamente en el Stage:

Tabla DWH	Tabla Stage	Descripción
Dim_Fecha	-	Maestro de fechas
Dim_AnalisisTiempo	-	Maestro para realizar análisis temporal (acumulados,

Tabla DWH	Tabla Stage	Descripción
		comparaciones, etc.)
Dim_Cliente	Custtable	Dimensión de clientes
Dim_Articulo	Inventtable	Dimensión de artículos
Dim_Proveedor	Vendtable	Dimensión de proveedores
Dim_Almacen	Inventlocation	Dimensión de almacenes
Dim_Analitica_CanalDistribucion Dim_Analitica_Comercial Dim_Analitica_Detalle Dim_Analitica_Naturaleza Dim_Analitica_Tipo	Dimensions	Dimensiones analíticas
Dim_Promocion	Promotion	Dimensión de promociones
Dim_Catalogo	Catalog	Dimensión de catálogos
Dim_Campaña	Campaing	Dimensión de campaña a nivel de fecha de la campaña anual
Fact_Catalogo_Articulos	Catalog	Hechos asociados a la relación existente entre hojas de

Tabla DWH	Tabla Stage	Descripción
		catálogos y los artículos que aparecen en ellas
Fact_Promociones_Segmentacion	Promotion	Hechos que permiten identificar los clientes objetivo de la segmentación de promociones
Fact_Ventas_Facturas	CustInvoiceJour CustInvoiceTrans	Hechos de las líneas de facturas de ventas
Fact_Ventas_Pedidos	SalesTable SalesLine	Hechos de las líneas de pedidos de ventas
Dim_Razones	Reason	Dimensión de tipificación de hechos atípicos de Ventas
Dim_Soluciones	Solution	Dimensión de Acciones de corrección o estrategia

Tabla DWH	Tabla Stage	Descripción
Cust_DAFO_Comercial	DafoComercial	Hechos de las decisiones en Clientes

5.2 Extracción, transformación y carga de datos

La arquitectura de la solución implementará un almacén de datos de ventas que incorporará las tablas de hechos y dimensiones necesarias para el análisis de las ventas.

El origen de los datos es la base de datos del ERP:

- Microsoft Dynamics AX 2009

Los datos se cargarán en el almacén utilizando la herramienta de extracción, transformación y carga de datos (ETL) **Microsoft SQL Server Integration Services**.

Se utilizará un modelado en estrella entre las tablas de hechos y las de dimensiones en el **almacén de datos** que se construirá en una base de datos relacional de SQL Server. Este modelado permite que las uniones entre las dimensiones y las tablas de hechos, no sean anidadas, ya que desde la tabla de hechos a cualquier dimensión sólo hay una unión.

Finalmente, a partir de las tablas de hechos y de dimensiones, se generará un cubo OLAP utilizando **Microsoft SQL Server Analysis Services**.

En la Figura 26 se puede ver gráficamente la representación del modelo híbrido propuesto.

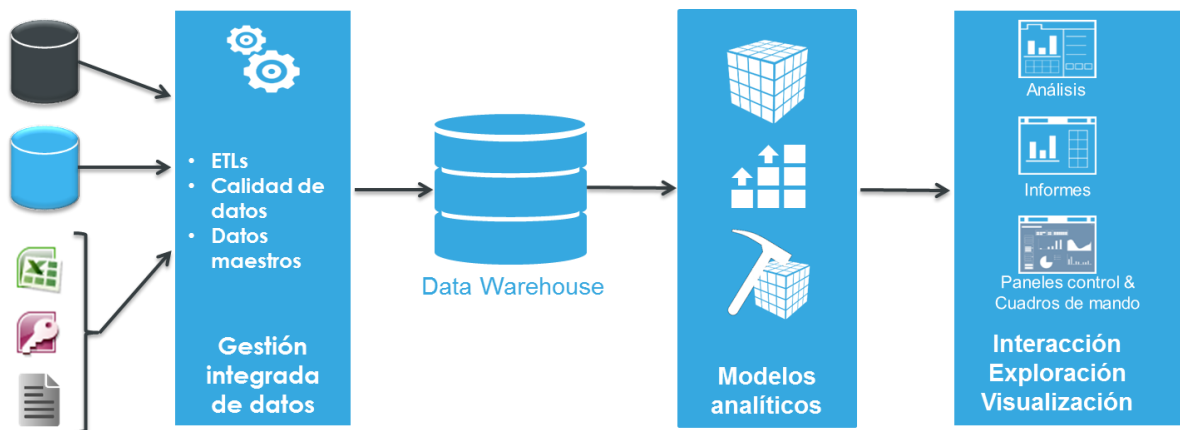


Figura 26: Modelo Híbrido

5.2.1 Estructura de BBDD

Las bases de datos que darán soporte a los distintos módulos de Business Intelligence se estructurarán en base a la Figura 27:

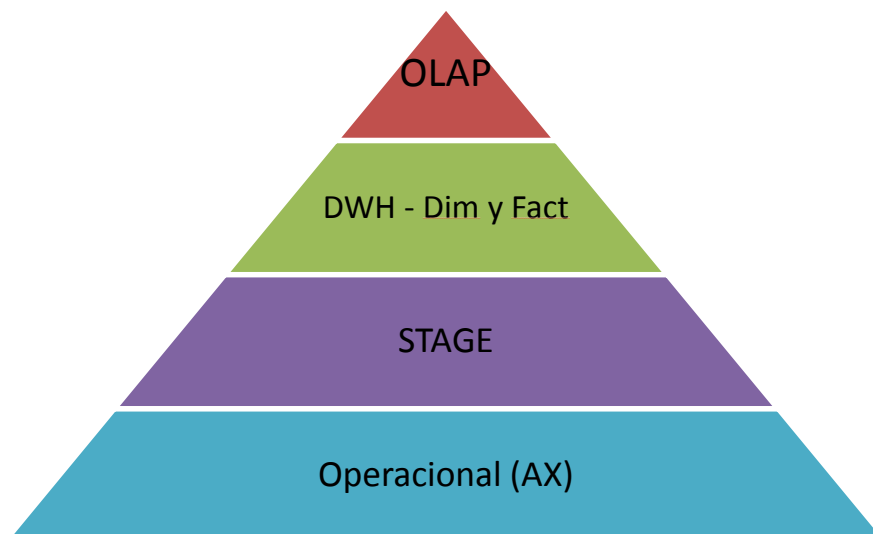


Figura 27: Fases 1,2 y 3

- **Stage:** Se replican las tablas y los campos necesarios de aquellos sistemas origen que interesen, como paso previo a la carga del DWH.

Todas las tablas son temporales, su información puede ser borrada en cualquier momento.

- **DWH – Dim y Fact:** en este esquema de BD se crean las tablas de dimensiones y las tablas de hechos, que dan soporte a la creación y carga de los cubos multidimensionales OLAP.
- **OLAP:** bases de datos multidimensionales orientadas al análisis de información y a la generación de informes y cuadros de mando.

Además, existe un esquema adicional:

- **Auxiliar:** esquema con tablas de gestión y control de los procesos de BI.

5.2.2 Carga de datos

La carga de datos se realizará en dos pasos:

1. **Carga Stage Área.** En esta primera carga se extraen los datos necesarios de los diferentes orígenes de datos sin transformaciones para almacenarlos en el área de recepción.
2. **Carga Data Warehouse.** En este paso se extraen y transforman los datos alojados en el área de recepción de información de modo que la información dentro de esta estructura es información consolidada válida para consultar.

La carga de maestros y dimensiones se hace incremental, insertando nuevos registros y actualizando aquellos que ya existen. En ningún momento se borran datos de estas tablas ya que las claves subrogadas deben permanecer invariables por registro.

La carga de transacciones y de hechos de ventas se realiza completa para un periodo determinado, es decir, se borran los registros del periodo en curso y se vuelven a cargar. Se deberá configurar como parámetro el periodo a recargar.

5.2.2.1 ETL Diario

Se ejecutarán cargas incrementales de forma periódica. Pendiente determinar el intervalo.

La carga de datos se realizará mediante 3 paquetes de Integration Services:

Carga de Stage

Carga de DWH

Procesado base de datos OLAP

5.2.2.2 Log de ejecución

En cada uno de los pasos del ETL diario se llama a un paquete dtsx de SSIS (Integration Services) que es el que tiene la lógica para la carga de datos en las tablas correspondientes.

En cada uno de estos paquetes, se controla la ejecución del mismo, escribiendo un log en una tabla auxiliar, con la siguiente información:

Campo	Descripción
Identificador	Auto numérico
TipoPaquete	Dimensión, Hechos, OLAP
CargaDescripcion	Nombre del paquete de SSIS ejecutado
ComputerName	Servidor desde el que se ejecuta
DTSx_Inicio	Fecha y hora de inicio de ejecución
DTSx_Fin	Fecha y hora de fin de ejecución
DTSx_CargaDesde	Fecha desde la que se realiza la carga para transacciones y hechos.
DTSx_CargaHasta	Fecha hasta la que se realiza la carga para transacciones y hechos.
LogStr	Comentarios de la ejecución.
Error	Si el paquete acaba con error, se activa este campo

5.3 Modelos analíticos (OLAP)

5.3.1 Dimensiones

A continuación se describen las dimensiones que se incorporan en el almacén de datos para el análisis de ventas.

5.3.1.1 Fecha

Dimensión de tipo tiempo a nivel de fecha. Esta dimensión es usada de manera habitual en proyectos de BI (Master Calendar), permitiendo al sistema enlazar las tablas de hechos con esta dimensión a través de claves numéricas, proporcionando al sistema enlaces (JOIN) más eficientes que los formatos de fecha.

Esta dimensión permite la generación de atributos personalizados relacionados con la fecha, permitiendo un crecimiento escalable de sus atributos, en el proyecto no se contempla la festividad del día, añadir un atributo festividad de tipo booleano permitiría usar este atributo como filtro o agrupación dentro del modelo analítico

Clave	Atributo	Descripción
PK	Fecha	Atributo con clave formada por el código de fecha en formato numérico y descripción la fecha
	Año	Año
	Trimestre	Atributo con clave formada por YYYYT y descripción del nombre del trimestre del año (2013 -1T).
	Mes	Atributo con clave formada por el identificador del mes en formato YYYYMM y descripción del nombre del mes del año (Enero 2013).
	Semana	Atributo con clave formada por YYYYWW y descripción del nombre de la semana del año (2013 –Semana 1).

	Nombre Mes	Nombre del mes, independiente del año (Enero, Febrero,...).
--	---------------	---

Se definen las siguientes jerarquías:

Por meses	Por semanas
Año	Año
Trimestre	Semana
Mes	Fecha
Fecha	

5.3.1.2 Análisis Tiempo

Dimensión especial que se utiliza para poder realizar cálculos sobre las medidas. Indica el momento temporal para el que se va a realizar el análisis, es decir, que a partir de un periodo de tiempo seleccionado se podrá analizar la información para ese periodo, para el mismo periodo del año anterior, el acumulado desde comienzo del año hasta ese periodo, año móvil, etc.

Por defecto está preseleccionado el miembro Actual. Es decir, que si no se filtra por esta dimensión, las medidas harán referencia a su valor durante el periodo de tiempo seleccionado.

Clave	Nombre	Descripción	Abrev.
1	Actual	Dato actual. Es el valor predeterminado de la dimensión.	Actual
2	Año Ant.	Igual periodo del año anterior	AA
3	Dif. Año Ant.	Diferencia en valor respecto al mismo periodo del año anterior	Dif_AA
4	Var. Año Ant	Diferencia en porcentaje respecto al mismo periodo del año anterior	Var_AA
5	Período Ant.	El periodo anterior al actual (mes anterior, trimestre anterior, etc.)	Ant

Clave	Nombre	Descripción	Abrev.
6	Dif. Per. Ant.	Diferencia en valor respecto al periodo anterior	Dif_Ant
7	Var. Per. Ant.	Diferencia en porcentaje respecto al periodo anterior	Var_Ant
8	Acum Anual	Acumulado anual desde el 1 de Enero	YTD
9	Acum Anual AA	Acumulado anual desde el 1 de Enero del mismo periodo del año anterior	YTD_AA
10	Dif. Acum AA	Diferencia en valor respecto al acumulado anual desde el 1 de Enero del mismo periodo del año anterior	Dif_YTD_AA
11	Var. Acum AA	Diferencia en porcentaje respecto al acumulado anual desde el 1 de Enero del mismo periodo del año anterior	Var_YTD_AA
12	Total Año Móvil	Total año móvil, son los 12 últimos meses	TAM
13	Total Año Móvil AA	Total Año móvil año anterior	TAM_AA

5.3.1.3 Cliente

Cada cliente tiene una cuenta de facturación asociada y además tiene una cuenta de agrupación que permitirá analizar las ventas para clientes de un mismo grupo.

Se definirán **3 dimensiones de cliente** asociadas a la facturación de ventas:

- Cliente que ordena la venta
- Cliente al que se le factura la venta
- Cliente grupo

Esta dimensión debe construirse de tipo lentamente cambiante en el tiempo, ya que interesa analizar el comportamiento del cliente en función de las personas que lo gestionan. Se generará un registro nuevo para cada cliente cada vez que en AX cambie el comercial, GAP (administración comercial) o agente (comercial externo) asignado incluyendo fechas de inicio y fin de la vigencia de ese registro.

Será necesario reconstruir el historial de personas asignadas a cada cliente a partir de la tabla de histórico para la carga inicial de la dimensión de cliente.

El resto de atributos del cliente se actualizarán normalmente en el almacén de datos cuando cambien en AX, no manteniendo su histórico de cambios.

Clave	Atributo	Descripción
PK	IdCliente	Es la clave de la dimensión. Id de la tabla de clientes.
	Cliente	Nombre del cliente.
	Código Cliente	Código del cliente en AX
Situación del cliente		
	Estado	Determina en qué situación se encuentra el cliente dentro del ciclo de vida de clientes. Se recalcula en AX periódicamente.
	Activo	Es un check que se utiliza para envíos comerciales de catálogos, etc. No tiene mucha calidad.
	Bloqueado	Indica si el cliente está bloqueado para ventas.
	Motivo del bloqueo	Indica el motivo por el que el cliente está bloqueado. Lista de valores.
Geografía del cliente		
	Localidad	Ciudad donde se encuentra el cliente
	Provincia	Provincia donde se encuentra el cliente
	CCAA	Comunidad autónoma donde se encuentra el cliente
	País	País donde se encuentra el cliente
	Código Postal	Código postal del cliente
Clasificación del cliente		
	Tipo cliente	Clasificación del cliente según tipologías.
	Canal distribución	Identifica el canal del cliente.

Clave	Atributo	Descripción
	Clase	Clasificación del cliente.
	Segmento	Segmentación del cliente de Marketing.
	Subsegmento	Segmentación del cliente de Marketing.
	Zona de ventas	Agrupación de comerciales
	Sector	Sector del cliente
	Potencial del cliente	Clasificación del cliente según lo que se espera de él para el futuro.
Dimensiones analíticas		
(se llevarán como dimensión a la venta, no visibles desde cliente)		
	Canal distribución	Dimensión analítica correspondiente al canal de distribución del cliente
	Empleado	Dimensión analítica correspondiente al comercial asociado al cliente
	Naturaleza	Dimensión analítica asociada al cliente
Gestión del cliente		
	Comercial actual	<p>Persona del departamento comercial que lleva la cuenta del cliente. Identifica la persona asignada en la actualidad al cliente.</p> <p>La clave es el usuario de directorio activo del comercial.</p> <p>Este campo es el que se utilizará para aplicar la seguridad de acceso a datos de clientes para el departamento comercial.</p>
(SCD)	Comercial histórico	<p>Persona del departamento comercial que llevaba la cuenta del cliente en un momento de tiempo concreto.</p> <p>Permite ver qué comercial estaba asignado en el momento de una venta a un cliente.</p>

Clave	Atributo	Descripción
		Atributo sujeto a cambios en el tiempo.
	Director comercial actual	<p>Persona responsable del comercial asignado a la cuenta del cliente en la actualidad.</p> <p>La clave es el usuario del directorio activo del director comercial.</p>
	GAP actual	Persona que mecaniza el pedido de venta. Identifica la persona asignada en la actualidad al cliente.
(SCD)	GAP histórico	<p>Persona que mecanizaba el pedido de venta del cliente en un momento de tiempo concreto.</p> <p>Atributo sujeto a cambios en el tiempo.</p>
	Agente actual	Persona comercial externo que trabaja con el cliente bajo comisión. Identifica la persona asignada en la actualidad al cliente.
(SCD)	Agente histórico	<p>Persona comercial externo que trabajaba con el cliente bajo comisión en un momento de tiempo concreto.</p> <p>Atributo sujeto a cambios en el tiempo.</p>
Otros		
	Grupo estadísticos	Grupo de segmentación comercial
	Modo entrega	Agencia de transporte por defecto
	Condiciones entrega	Condiciones de entrega por defecto del cliente
	Descuento línea	Grupo de descuento de línea
	Descuento multilínea	Grupo de descuento multilínea

Clave	Atributo	Descripción
	Grupo artículo adicionales	Grupo de Artículos adicionales del cliente
	Forma de pago	Forma de pago por defecto del cliente
	Condiciones de pago	Condiciones del pago del cliente
	Recepción de pedidos	Canal de transmisión del pedido del cliente (mail, teléfono, internet, Edi ...)
	Política stock	Sin Almacén, Almacén, Plataforma, Operador Logístico
Gestión registros dimensión lentamente cambiante		
	Fecha inicio registro	Fecha desde la que es efectivo el registro de cliente.
	Fecha fin registro	Fecha hasta la que es efectivo el registro de cliente.
	Versión actual	Check para indicar si es el registro vigente en la actualidad.

(SCD) – Atributos de dimensión lentamente cambiante/ slowly changing dimension.

Se define la siguiente jerarquía para la dimensión cliente:

Geografía
País
CCAA
Provincia
Localidad
Cliente

5.3.1.4 Artículo

Son los productos que se venden en la empresa

Esta dimensión debe construirse de tipo lentamente cambiante en el tiempo, ya que interesa analizar el comportamiento del artículo en sus distintos estados del ciclo de vida. Se generará un registro nuevo para cada artículo cada vez que en AX cambie su estado, incluyendo fechas de inicio y fin de la vigencia de ese estado. Estos estados contemplan el ciclo de vida clásico de un artículo (Ninguno, Provisional, Novedad, Activo, Sin Continuidad, Chanca y Baja)

El resto de atributos del artículo se actualizarán normalmente en el almacén de datos cuando cambien en AX, no manteniendo su histórico de cambios.

Será necesario reconstruir el historial de estados de cada artículo en el tiempo a partir del log de seguimiento a cambios de AX.

Clave	Atributo	Descripción
PK	IdArticulo	Es la clave de la dimensión. Id de la tabla de artículos.
	Artículo	Nombre del producto.
	Código artículo	Código del producto
	Estado	Situación del artículo. Atributo sujeto a cambios en el tiempo.
Clasificación de productos		
	Tipología	Clasificación del artículo
	Tipo de artículo	Tipo estándar de producto según AX
	Gran familia	Nivel 1 de clasificación de productos. El artículo tendrá asignada una subfamilia y cada subfamilia pertenece a una gran familia.

Clave	Atributo	Descripción
	Subfamilia	Nivel 2 de clasificación de productos.
	Campaña	Un artículo puede estar asignado a una campaña temporal. Cada campaña tendrá un intervalo de fechas anuales en las que se ejecuta. Permitirá asociar las ventas a una dimensión campaña.
	Grupo compradores	Persona responsable de la compra del artículo.
	Lote	A partir de la configuración del grupo de dimensiones de inventario, se realiza un cálculo para indicar si el artículo requiere lote o no.
	Intrastat	Clasificación oficial de artículos en categorías.
Dimensiones contables analíticas (se llevarán como dimensión a la venta, no visibles desde artículo)		
	Canal distribución	Dimensión analítica contable 1
	Detalle	Dimensión analítica contable 2
	Tipo	Dimensión analítica contable 3
Royaltie (Los artículos asociados a marcas comerciales tienen que pagar un porcentaje a las marcas)		
	Tipo de royaltie	Royaltie aplicado al artículo
	Intervalo porcentaje royaltie	Intervalo de porcentajes para royaltie

Clave	Atributo	Descripción
	Descripción royaltie	A partir de los dos campos anteriores se obtiene la descripción y el porcentaje a aplicar en la venta.
	% Royaltie	Porcentaje que se aplica sobre el importe de venta neto para calcular el royaltie a pagar a la marca.
Otros		
	Grupo impuestos artículo	Indica que tipo de impuestos se aplica al artículo
	Identificador franquicia	Todas las licencias, disponen de franquicias que identifican principalmente al personaje
	Descuento línea	Descuento de línea de AX
	Descuento multilínea	Descuento de multilínea de AX
	Descuento comercial máximo	Descuento máximo permitido
	%Portes	Porcentaje de portes por unidad de venta
Gestión registros dimensión lentamente cambiante		
	Fecha inicio registro	Fecha desde la que es efectivo el registro del artículo.
	Fecha fin registro	Fecha hasta la que es efectivo el registro del artículo.
	Versión actual	Check para indicar si es el registro vigente en la actualidad.

Se define la siguiente jerarquía:

Clasificación
Gran familia
Subfamilia
Artículo

5.3.1.5 Proveedor

Son los distintos proveedores de productos.

Cada artículo está asociado a un único proveedor por lo que las ventas a cliente final también podrán analizarse a nivel de proveedor.

Para el histórico de ventas, se asumirá que el proveedor actual es el proveedor de la venta.

Clave	Atributo	Descripción
PK	Proveedor	Nombre del proveedor. Su clave es el id de la tabla de proveedores.
	Código proveedor	Código del proveedor en DEKORA
	Responsable proveedor	Persona responsable del proveedor de DEKORA.

5.3.1.6 Campaña

Son las campañas temporales que agrupan artículos y que se repiten cada año.

Esta dimensión permitirá comparar ventas realizadas dentro de la misma campaña en años diferentes. Para asociar las ventas a las campañas, se generará una nueva tabla en AX donde se defina para cada campaña y año los intervalos de fechas en los que se lleva a cabo.

Cada artículo está asociado como máximo a una campaña por lo que las ventas a cliente final también podrán analizarse a nivel de campaña según la fecha de la venta. Sólo se asociarán a campaña las ventas de artículos que tienen definida una campaña para el intervalo de fechas indicado en la campaña anual. El resto de ventas de otros artículos o fuera del intervalo de fechas de las campañas irán a campaña Indefinida.

En la dimensión de campaña tendremos un registro por cada fecha del año incluida en una campaña.

Clave	Atributo	Descripción
PK	Fecha campaña	Fecha incluida en la campaña. En una misma fecha nunca se pueden solapar dos campañas distintas. La clave será el id numérico correspondiente a la fecha en formato YYYYMMDD.
	Campaña	Nombre de la campaña
	Año campaña	Año al que corresponde la campaña. Por ejemplo la campaña de Navidad del año 2014 incluirá fechas desde diciembre de 2014 a enero de 2015.
	Número día campaña	Número secuencial del día al que corresponde la fecha para esa campaña anual.

5.3.1.7 Catálogo

Cada artículo está asociado a varios catálogos simultáneamente por lo que la venta no podrá vincularse directamente a un catálogo de productos por fecha.

Se creará una relación **muchos a muchos** entre catálogo y ventas a través de un grupo de medidas intermedio formado por la relación de catálogo y artículo.

Esto permitirá utilizar los catálogos como filtros para seleccionar varios artículos pertenecientes a un mismo catálogo.

Interesa también trabajar con las hojas del catálogo donde está el artículo por lo que esta dimensión bajará hasta la hoja del catálogo.

Clave	Atributo	Descripción
PK	Hoja catálogo	Página del catálogo donde aparece el artículo. Su clave será un Id auto numérico generado por cada código de catálogo y hoja diferente.
	Catálogo	Nombre del catálogo.
	Hoja	Número de página.
	Fecha catálogo	Fecha en la que se publicó el catálogo.

5.3.1.8 Promociones

Marketing genera promociones que aplican a segmentos de clientes. Interesa poder analizar qué clientes de esos segmentos a los que aplica una promoción no han aprovechado la promoción. Para ello se generará un grupo de medidas que relacionará los clientes con las promociones que le aplican.

Las ventas que se realizan en promoción se relacionarán con una nueva dimensión promoción.

Clave	Atributo	Descripción
PK	Promoción	Nombre de la promoción. Su clave será el Id de la promoción.
	Código promoción	Código de la promoción en el ERP
	Modo	Forma en la que se aplica la promoción
	Tipo	Clasificación de la promoción

Clave	Atributo	Descripción
	Fecha inicio	Fecha de comienzo de la promoción
	Fecha fin	Fecha de finalización de la promoción

5.3.1.9 Almacén

Las ventas también están asociadas al almacén del que sale la mercancía.

Distinguiendo, almacenes físicos, almacenes de tránsito y almacenes externos.

Clave	Atributo	Descripción
PK	Almacén	Nombre del almacén.
	Código almacén	Código del almacén en el ERP
	Tipo	Clasificación del almacén (físico, tránsito, externo)
	Estado	Situación del almacén

5.3.1.10 Dimensiones analíticas

Se generarán 5 dimensiones analíticas contables asociadas a las ventas:

- **Canal distribución:** es el canal que tiene asociado el cliente en el momento de la venta.
- **Comercial:** es el empleado que tiene asignada la cuenta como comercial en el momento de la venta.
- **Naturaleza:** dimensión analítica que depende del cliente de la venta.
- **Detalle:** dimensión analítica que depende del artículo de la venta.
- **Tipo:** dimensión analítica que depende del artículo de la venta.

5.3.1.11 Abono y residuo

Dimensión que permitirá analizar los pedidos devueltos por los clientes en función de si se puede aprovechar la devolución o no.

Incluirá 3 opciones:

- **Abono:** devolución del cliente que vuelve al almacén.
- **Abono y residuo:** devolución del cliente que no puede reutilizarse y no vuelve a almacén.
- **Resto ventas:** venta correcta sin devolución.

El cálculo del coste en facturas de ventas será especial en el caso de abono y residuo, ya que no se recuperan costes. El importe de la venta es negativo y el coste es cero.

5.3.2 Grupos de medidas

5.3.2.1 Facturas ventas

Las medidas que surgen directamente de la tabla de hechos son las siguientes:

Medida	Descripción	Agregación
Importe ventas	Importe de la venta	Suma
Descuento	Importe descontado por promoción	Suma
Precio tarifa	Importe de la venta antes de aplicar el descuento (Importe ventas + Descuento)	Suma
Importe royaltie	Importe que se obtiene de aplicar el porcentaje de royaltie que tiene el artículo sobre el importe de la venta	Suma
Costes	Coste de la venta.	Suma

	Se calcula a partir del precio del artículo. No se utilizan las transacciones de inventario ya que no se realizan cierres de inventario en AX.	
Cantidad (inventario)	Cantidad asociada a la venta en la unidad de inventario	Suma
Cantidad (consumo)	Cantidad asociada a la venta en la unidad de consumo	Suma

Se añaden los siguientes contadores:

Medida	Descripción	Agregación
Nº Clientes ventas	Contador de clientes distintos que compran.	Contador distinto
Nº Artículos ventas	Contador de artículos distintos que compran.	Contador distinto

Se añaden los siguientes valores calculados:

Cálculo	Descripción
Aportación	Importe ventas – Costes – Importe royaltie
Rentabilidad	Beneficio / Importe ventas

5.3.2.2 Pedidos ventas

Se generarán hechos de pedidos de ventas con el fin de analizar la calidad del servicio ofrecido a los clientes.

Medida	Descripción	Agregación
Importe pedido	Importe del pedido.	Suma
Plazo de entrega	Días que transcurren desde que se genera el pedido hasta que se completa.	Suma
Número de líneas	Contador de líneas del pedido.	Contador
Cantidad pedida	Cantidad solicitada por el cliente en unidades de inventario	Suma
Cantidad servida	Cantidad solicitada por el cliente que se ha servido en unidades de inventario.	Suma
Cantidad no servida	Cantidad solicitada por el cliente que no ha podido ser servida en unidades de inventario.	Suma

5.3.2.3 Segmentación promociones

Este grupo de medidas permitirá relacionar las distintas promociones con los segmentos de clientes a los que aplica y con los clientes que finalmente se han beneficiado de la promoción.

En esta tabla se generará un registro por cada promoción y cliente a los que se aplica.

Medida	Descripción	Agregación
Nº clientes aplica promoción	Contador de clientes distintos en la tabla.	Contador distinto

5.3.2.4 Catálogos de artículos

Grupo de medidas para relacionar los artículos con los distintos catálogos en los que aparece. Servirá como grupo de medidas intermedio para la relación muchos a muchos entre artículos y hojas de catálogo.

Medida	Descripción	Agregación
Nº artículos catálogo	Contador de registros.	Contador

5.3.3 Esquema estrella

Todas las dimensiones y medidas explicadas anteriormente, serán implantadas sobre distintos esquemas estrellas.

5.3.3.1 Facturas ventas

Este esquema permitirá en análisis por parte de los usuarios de las medidas obtenidas de las facturas de ventas y relacionadas, desde todas las dimensiones relacionadas, en la Figura 28 se puede ver la representación gráfica del esquema estrella de Facturas de Ventas.

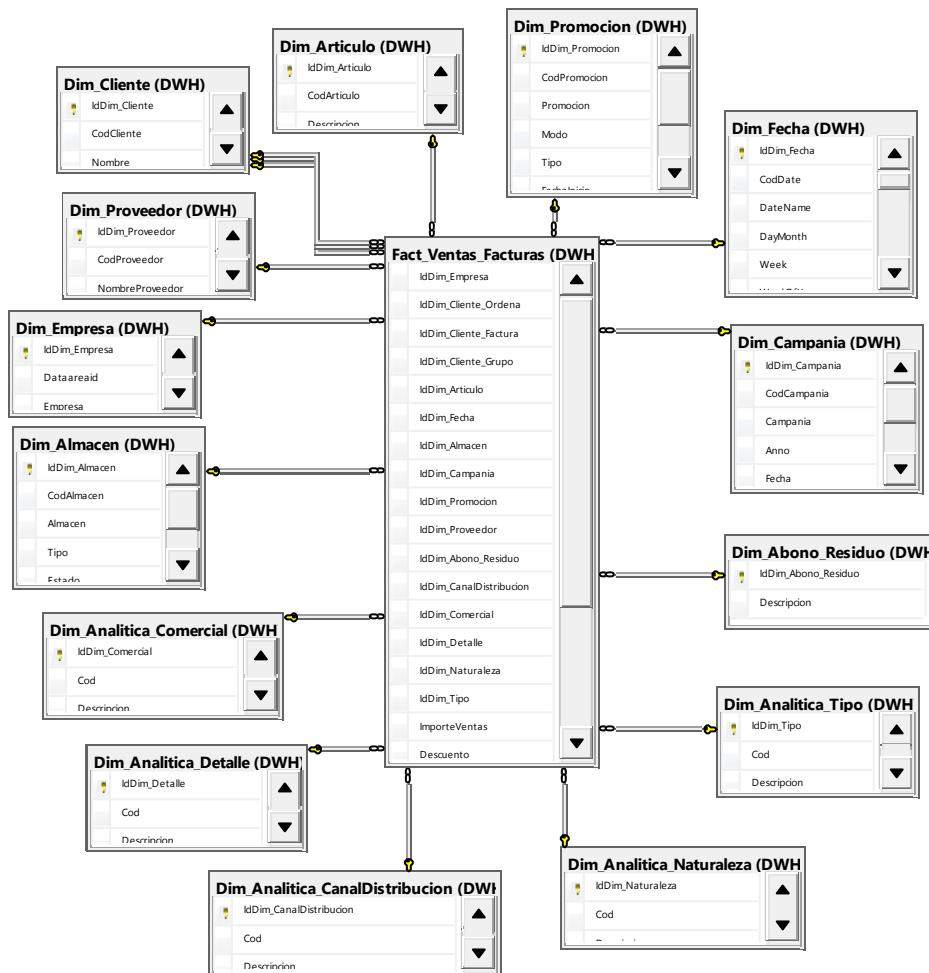


Figura 28: Esquema Estrella Facturas de Venta

5.3.3.2 Pedidos ventas

Este esquema permitirá en análisis por parte de los usuarios de las medidas obtenidas de los pedidos de ventas y relacionadas, desde todas las dimensiones relacionadas, en la Figura 29 se puede ver la representación gráfica del esquema estrella de Pedidos de Ventas.

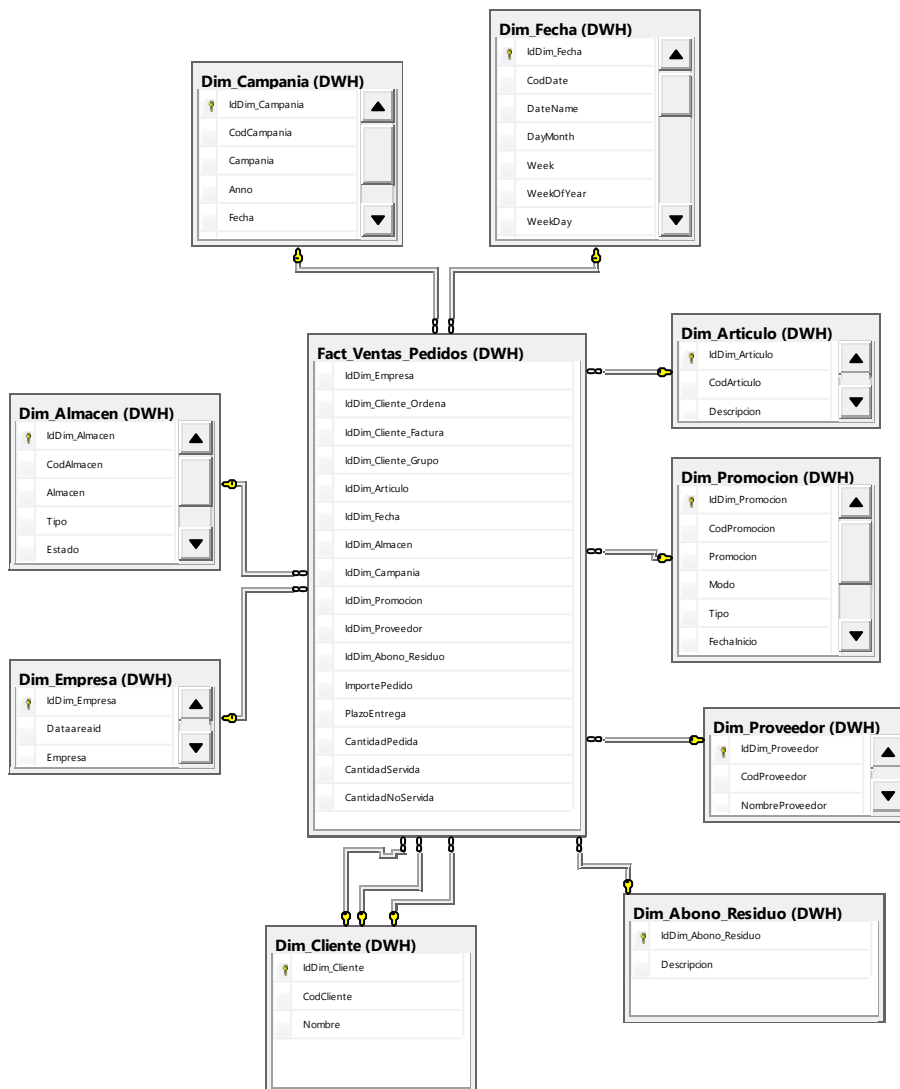


Figura 29: Modelo Estrella Pedidos de Venta

5.3.3.3 Segmentación promociones

Este esquema permitirá en análisis por parte de los usuarios de las medidas obtenidas de las promociones de marketing y relacionadas, desde todas las dimensiones relacionadas, en la Figura 30 se puede ver la representación gráfica del esquema estrella de Promociones de Marketing.

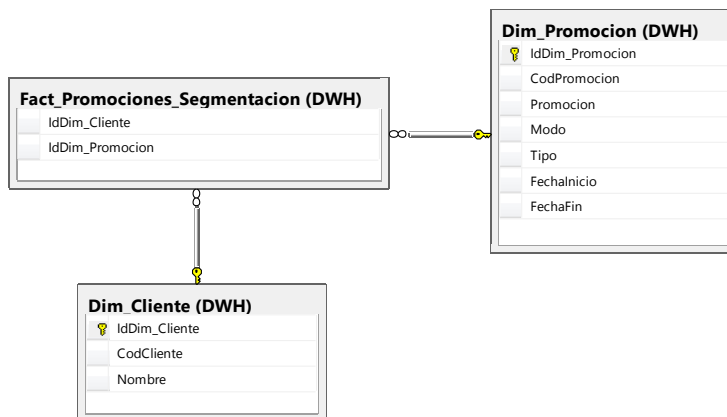


Figura 30: Modelo Estrella Promociones

5.3.3.4 Catálogos de artículos

Este esquema permitirá en análisis por parte de los usuarios de las medidas obtenidas de los catálogos de artículo y relacionadas, desde todas las dimensiones relacionadas. Es importante resaltar que el objetivo de este esquema consiste en valorar, el desarrollo de marketing de catálogos, incluir medidas de otros esquemas tiene riesgos inherentes debido principalmente a que el mismo artículo, puede estar incluido en varios catálogos simultáneamente (catálogos con validez temporal similar), esto supone que si se analiza la facturación de un artículo en un periodo determinado agrupando por catálogo, la facturación se duplica en todos los catálogos activos a los que pertenece ese artículo. En la Figura 31 se puede ver la representación gráfica del esquema estrella de Catálogos de Artículo.

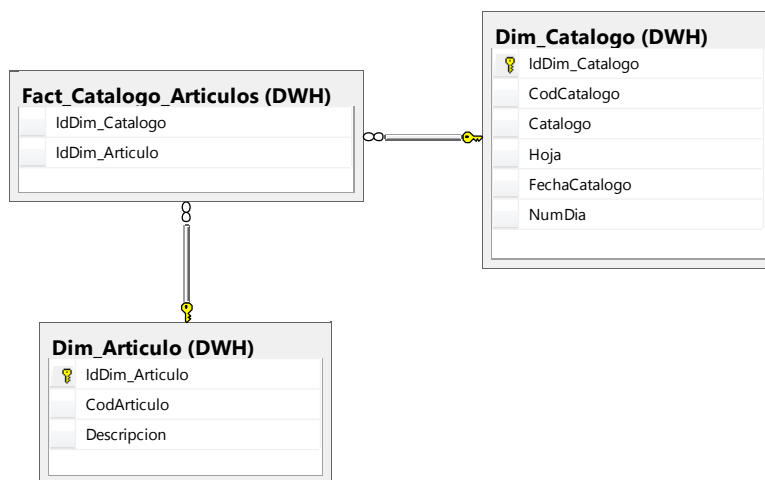


Figura 31: Modelo Estrella Catálogos de Artículo

5.3.4 Roles seguridad

Se definen 5 roles para gestionar la seguridad de acceso a la información:

Rol	Descripción
Administrador	Rol para el administrador de la BBDD OLAP de ventas.
Dirección general	Rol para acceso completo al cubo de ventas. No hay limitación por dimensión ni medidas. El usuario que accede a esta BBDD con este rol puede ver todo.
Marketing	Rol para acceso completo al cubo de ventas.
Director comercial	Rol para acceso al cubo de ventas por parte de los directores comerciales. Estos usuarios podrán acceder a los datos de su cartera actual de clientes y de las carteras de los comerciales que estén a su cargo.

	Se filtrará la dimensión Cliente por el campo Director comercial asignado utilizando el usuario que se conecta.
Comercial	<p>Rol para acceso al cubo de ventas por parte de los comerciales. Estos usuarios podrán acceder a los datos de su cartera actual de clientes.</p> <p>Se filtrará la dimensión Cliente por el campo Comercial asignado utilizando el usuario que se conecta.</p>

5.4 Informes y cuadros de mando

En este punto se realiza la explotación de los datos y son los usuarios de negocio, los trabajadores de la información. Estos trabajadores están sementados por su necesidad de información, y es necesario preguntarse, cuáles son sus necesidades analíticas.

Las verdaderas herramientas de descubrimiento de datos son mucho más raras, ya que son mucho más difíciles de desarrollar. Como dijo una vez Steve Jobs, simplicidad real es muy difícil.

De esta manera, podemos dividir los trabajadores de la información en tres grupos:

1. Jefes
2. Analistas
3. Estadistas

Esta división es especialmente interesante porque existe una correspondencia bastante exacta con las herramientas disponibles en el mercado:

1. Cuadros de mando

2. Reporting & OLAP
3. Data mining

5.4.1 Analistas

Estos usuarios de negocio plantean preguntas concretas al modelo OLAP, y generan informes en base a las necesidades, dentro del modelo propuesto, los analistas son los comerciales y los agentes de marketing, que en base a su visión del negocio bucean en los datos para preguntarse cosas concretas. Estas preguntas a su vez generan informes que ellos mismos construyen en base al modelo analítico, resolviendo, que compran mis clientes, como se está desarrollando la campaña, son rentables mis clientes, se venden todos los artículos de mi catálogo, he de potenciar familias de artículos, como se están comportando las distintas licencias de productos...

5.4.2 Informes funcionales

Estos informes son el resultado de la necesidad de información de los analistas, pretender dar respuesta a realidades concretas, estas realidades tienen que ser realidades representativas de la estrategia de la empresa

5.4.2.1 Ventas por Familia

En este informe se analiza las ventas por familias de productos, se crea una jerarquía (Falso MOLAP), Gran Familia → Familia → Sub Familia → Artículo, ofreciendo las siguientes métricas: Ventas, Aportación, Rentabilidad, Nº de Clientes (distintos clientes), cantidad (unidades de inventario vendidas). En la Figura 32 se muestra el informe generado



Figura 32: Ventas por Familia

5.4.2.2 Ventas por Campaña

En este informe se analizan las ventas de productos de campaña en las distintas campañas, ofreciendo al usuario Campañas→ Meses→ Clientes→ Artículos, en este caso las métricas son: Ventas, clientes, Portfolio (Nº de artículos distintos), analizando las variaciones de estas en un periodo de 3 años.

Particularmente este informe utiliza la dimensión fecha de campaña, la cual puede abarcar un periodo que incluya dos años distintos, ejemplo San Valentín 2015 empieza su venta el 1 de diciembre de 2014 y finaliza en 31 de Marzo de 2015 (si bien se podría suponer que San Valentín finaliza el 14 de Febrero, las devoluciones son parte de esta campaña y tiene ese límite máximo de devolución). En la Figura 33 se muestra el informe generado.

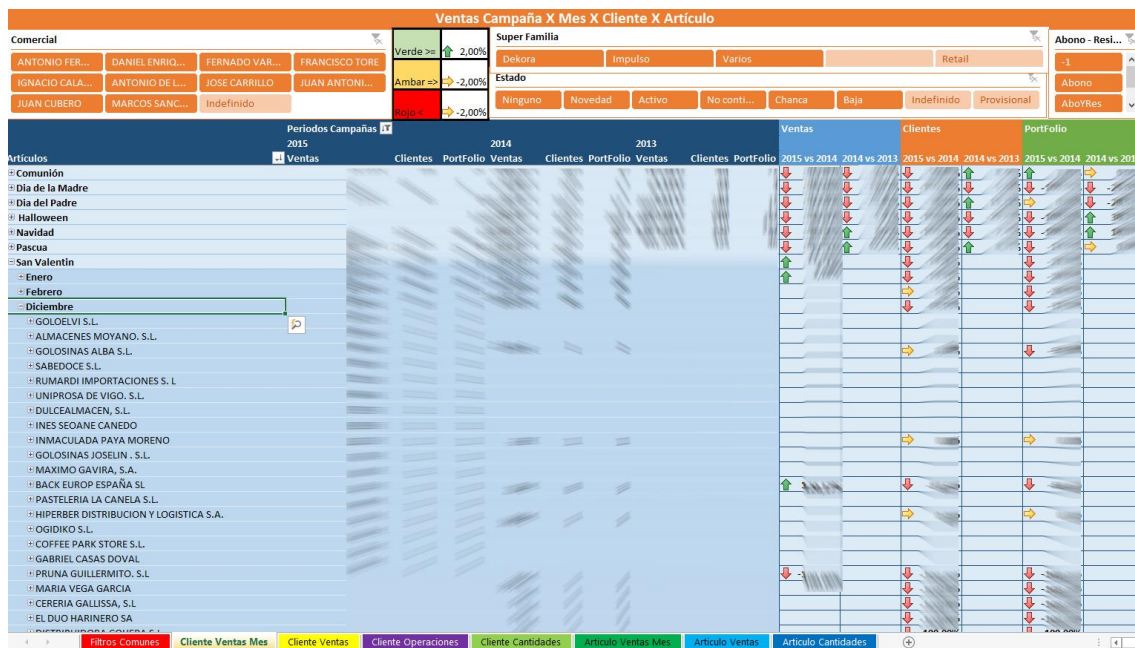


Figura 33: Ventas por Campaña

5.4.2.3 Ventas por Franquicia

En este informe se trata de analizar como las distintas franquicias (personajes de licencias concretas) evolucionan sus ventas, clientes y referencias, ofreciendo Franquicia→Artículos, en este caso las métricas son: Ventas, clientes, Portfolio (Nº de artículos distintos), analizando las variaciones de estas en un periodo de 3 años. En la Figura 34 se muestra el informe generado.



Figura 34: Ventas por Franquicia

5.4.2.4 Ficha de Cliente

Este informe opera directamente contra la dimensión cliente como partida de sus operaciones, y ofrece al usuario La ficha funciona por:


1. Comercial: Un comercial puede acceder a los clientes que tiene asignados exclusivamente
2. Zona: 2 Zonas (los jefes comerciales pueden ver las fichas de cualquier cliente de su zona)
3. Dirección/Marketing: Cualquier cuenta

Este documento contiene los siguientes elementos

Cabecera, en la Figura 35 se muestran los datos de cabecera

Ciente Pedido		TAM (12 Últimos meses)	Internet
Año		Pagos	
Pedido Medio		Días	Servicio

Figura 35: Cabecera Ficha de Cliente

1. Editable: Cliente Pedido/ Cliente Grupo: es un la lista de selección, para poder analizar una cuenta de manera individual, o de manera conjunta con otros clientes ejemplo Carrefour, esta opción se puede usar para agrupar clientes, pero no sólo tipo Carrefour, Back Europe, La Zentral ... si tenéis clientes que cambian de CIF, crear un grupo con la nueva cuenta, de tal forma, que ambos clientes funcionaran como uno sólo
2. Editable: Cuenta: Poner el número de cliente a analizar (el sistema solo mostrará los resultados si el cliente es vuestro, jefes comerciales todas las cuentas de los clientes de su zona, dirección/marketing todas las cuentas)
3. Botón: La lupa: Recalcula todos los datos de la ficha en base a la cuenta
4. Botón: Interrogación: Como no todo el mundo conoce el número de sus clientes he creado un formulario de búsqueda por nombre de cliente, filtrar y click sobre el icono . En la Figura 36 se muestra el formulario de búsqueda generado.

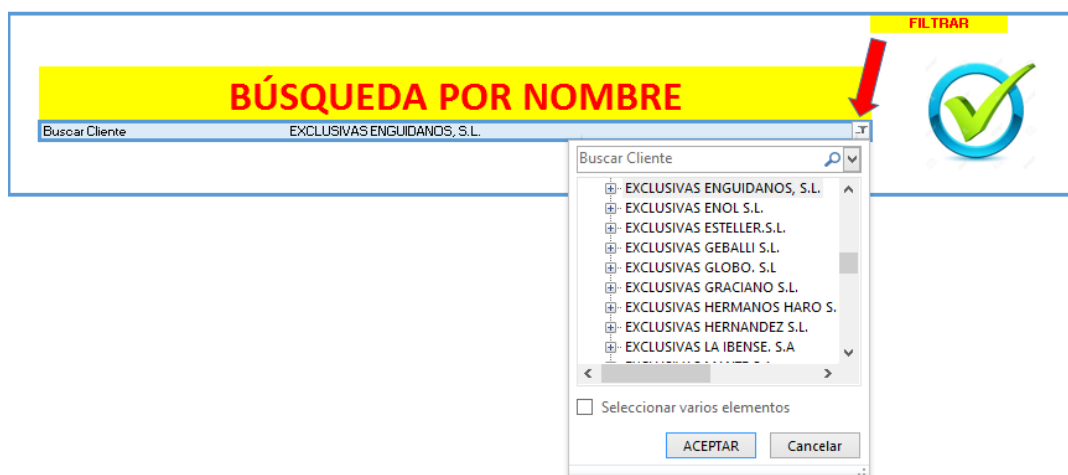


Figura 36: Búsqueda por Nombre de Cliente

5. Año: cambiando este valor a un año concreto el sistema calculará la última fecha pasada disponible en ese año (en 2015 sería ayer, mientras que en otros años sería el 31-12-XXXX)
6. Informativo: Pedido Medio: Calcula el valor medio de los Pedidos de un cliente
7. Informativo: Pagos: Realizados durante los 12 últimos meses
8. Informativo: Días: Valor medio de pago realizado por Euro en los 12 últimos meses

9. Informativo: Manual/Internet/Edi: Ventas realizadas por cualquiera de esos medios

En la Figura 37 se muestra el informe completo de la Ficha de Cliente.



Figura 37: Ficha de Cliente

5.4.3 Cuadro de mando (CMI)

5.4.3.1 Cuadro General

Este cuadro está diseñado con objetivo de usuario de negocio de alto nivel pueda ver el estado, de los principales indicadores de la empresa, en base a los roles de seguridad, cada usuario del sistema recupera un cuadro con la información que le concierne, por lo que el director general vería todo, el jefe comercial los datos generados por su zona y finalmente el comercial, el cual solo puede ver los datos de sus ventas.

Este documento contiene los siguientes elementos referidos siempre al periodo actual hasta el mismo periodo del año anterior, generalmente se utiliza la fecha del día anterior para la generación de los periodos de comparación.

1. Indicadores de Cabecera

- a. Ventas totales, acompañadas por su diferencia y su variación
- b. Aportación, acompañadas por su diferencia y su variación
- c. Ventas por zonas Comerciales , acompañadas por su diferencia y su variación
- d. Cobros TAM (Tasa Anual Media) y Días medios de Cobro TAM
- e. Ventas y Aportación TAM (Tasa Anual Media) y Días medios de Cobro TAM
- f. Ventas por zonas grandes familias , acompañadas por su diferencia y su variación
- g. Var Ventas
- h. Var Rentabilidad
- i. Var Descuento
- j. Var Aportación

2. Bloque 1

- a. Ventas y Rentabilidad por Canal
- b. Gráfico de Tarta de las ventas por canal
- c. Repartición de líneas y clientes por canal de entrada (como se introduce en el sistema), EDI [\[32\]](#) Electronic Data Interchange, Internet o Manual
- d. Gráfico de Tarta de las líneas por mecanismo de introducción al sistema

3. Bloque 2

- a. Ventas, Aportación y Rentabilidad por Zona y Comercial (periodo actual, periodo anterior, diferencia y variación)

4. Bloque 3

- Ventas Facturas/Abonos/Abonos no recuperables
- Ventas por Tipo de Cliente
- Ventas, Aportación y Rentabilidad por Gran Familia de productos, mostrando periodo actual, Diferencias y Variaciones

En la Figura 38 se muestra el Cuadro de Mandos Generado para el control de las ventas.

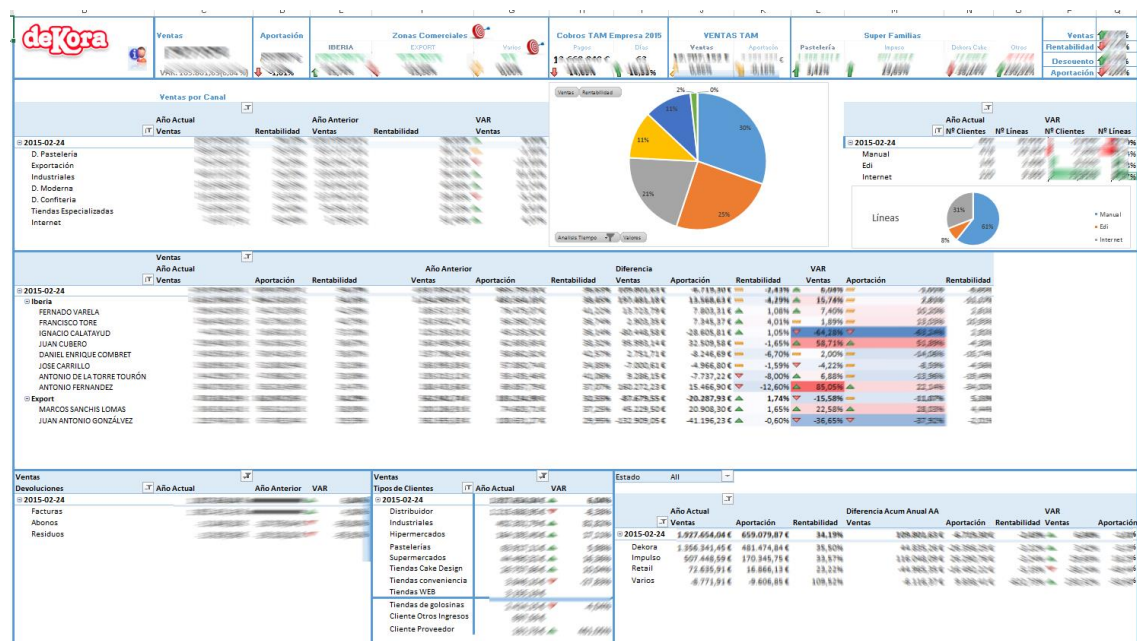


Figura 38: CMI Ventas

5.4.3.2 Cuadro Reducido

Con este cuadro, se aporta la cantidad de información mínima para el usuario de negocio, a través del cuadro general se obtiene una versión reducida, para dar soporte a la movilidad. En la Figura 39 se muestra el cuadro de ventas reducido.


	Ventas		Aportación	
				
				
Zonas Comerciales				
				
TAM				
				
Super Familias				
Dekora	Impuso		Dekora Cake	
				
Ventas				
Rentabilidad				
Descuento				
Aportación				

Figura 39: CMI Ventas

5.4.3.3 Movilidad

El CMI permite el auto envío programado o manual de los cuadros de la empresa a direcciones preestablecidas en formatos jpg o pdf, permitiendo al usuario de negocio recibir el cuadro referido siempre a su segmento de negocio

1. Director General y Marketing
2. Directores comerciales Iberia y Director General

3. Director comercial Exportación

El interface mostrado en la Figura 40 permite el envío por mail del cuadro general o reducido (en formato de imagen JPG o PDF, Acrobat Reader).

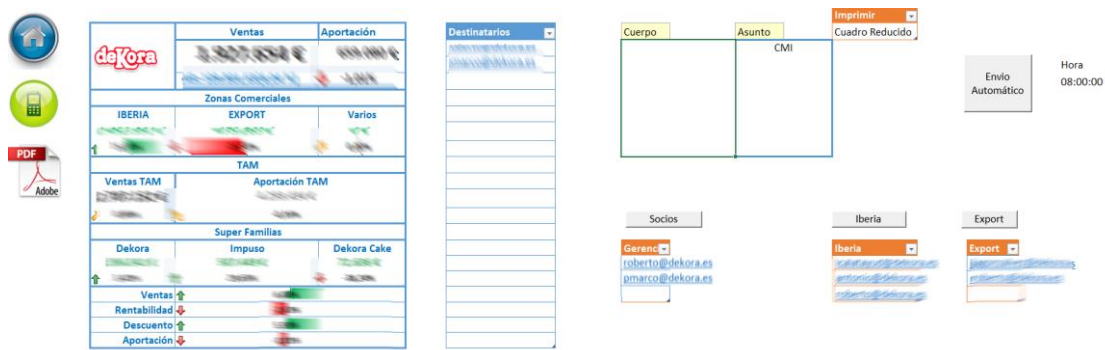


Figura 40: Movilidad CMI

5.5 Retroalimentación de Sistema

Como parte distintiva del proyecto, se pretende que la información aportada por el modelo analítico desarrollado, retroalimente al sistema de BI, para tal efecto se creó en los modelos analíticos. Encontramos cuatro partes diferenciadas claramente:

1. Análisis de la información para la toma de decisiones
2. Identificación y clasificación de situaciones atípicas mediante la estrategia DAFO
3. Toma de decisiones
4. Retroalimentación del sistema (primero al transaccional y posteriormente al BI)
5. Análisis de resultados.

5.5.1 Análisis de la información para la toma de decisiones

Este punto se ha basado principalmente en las unidades de medidas alineadas con el negocio, la empresa necesita ventas, rentabilidad, aportación, clientes, líneas, incrementos de las ventas.

El sistema proporciona las diferencias en las unidades de medidas acaecidas en periodos concretos, en este caso, anuales. El comercial analiza las diferencias en ventas, aportación, rentabilidad, Número de artículos, Número de líneas, y determina los valores atípicos, tanto los negativos como los positivos, Debilidades y Fortalezas. Cada medida aporta un valor estratégico para la compañía y analizados de manera conjunta determinan el estado de salud de un cliente, desde el punto de vista de la estrategia comercial.

5.5.2 Identificación y clasificación DAFO

De cara a realizar una identificación alineada con los objetivos de la empresa, se clasifican los valores extremos y se establece la relación con las distintas potencialidades, limitaciones, riesgos y desafíos que la empresa ha considerado como representativos en el plan estratégico. En casos de no poder identificarlo, el valor quedará como indeterminado para su posterior estudio. En la Figura 41 se puede ver el informe DAFO utilizado para los clientes.

DAFO Clientes									
<div>Filtros</div>		<div>Reportar</div>		Verde >=	30,00%				
				Ambar =>	20,00%				
				Rojo <	20,00%				
Diferencias									
Cientes	Ventas	Aportación	Rentabilidad	Nº Artículos	Nº de Clientes	Nº Líneas	VAR	Ventas	
2014-12-31									
DULMONT PRODUCTOS DE OBRADOR SL									
REPRESENTACIONES LLUCH. S.L.									
ATAVI S.A.									
SABATE DOLÇ I SALAT S.L.									
DIEXCA S.A.									
COMPLEMENTOS PAST MARTI SA									
CARTONAJES JUMARVA SA									

Figura 41: Análisis DAFO Cliente

5.5.3 Toma de Decisiones

Una vez determinada la matriz DAFO, se buscará las distintas estrategias diseñadas para hacer frente al factor combinado (potencialidades, limitaciones, riesgos y desafíos).

Estas estrategias, proporcionarán información contrastada, indicando en todo momento el nivel de éxito obtenido en decisiones anteriores.

Finalmente el usuario tomará la decisión que más conveniente para afrontar el factor combinado.

Sirva como ejemplo la clasificación DAFO realizada por un comercial de sus clientes que más han descendido en ventas:

1. Filtrar Variación de la Ventas en que se haya descendido más de un 20%
2. Filtrar descensos superiores a 1.000 €
3. Generar matriz DAFO por cliente
4. Identificar para las bajadas si trata de limitaciones, riesgos o desafíos (combinaciones que incluyes amenazas, debilidades o ambas).
5. Buscar estrategias anteriores orientadas para hacer frente a factor combinado.
6. Analizar las estrategias más afines al problema en cuestión, contextualizar el problema con las distintas estrategias. Valorando su grado de éxito en decisiones anteriores.
7. Determinar una o varias estrategias a implementar, determinado el espacio temporal de la estrategia y su objetivo cuantitativo, por ejemplo aplicar un descuento adicional del 5% desde el día 01/06/2015 hasta el 31/09/2015 se espera pasar de una variación actual de -1500 € a -500 €.

5.5.4 Fichero de Retroalimentación

La muestra Obtenida ya clasificada se trasladará a una tabla Excel donde se indicará.

1. Fecha de Exportación
2. Empleado
3. Factor Combinado
4. Tipo de Factor Combinado
5. Estrategia a Implementar
6. Valor de la estrategia.
7. Cuenta de Cliente
8. Diferencia en Venta Actual
9. Diferencia en Venta Objetivo
10. Fecha de Inicio
11. Fecha de Fin
12. Observaciones

En este punto el comercial genera un fichero que se importará en el ERP, la hoja en cuestión se guarda en formato CSV y a través de un importador, se vuelca nuevamente en el ERP. En la Figura 42 se puede ver el informe para la exportación comercial de la toma de decisiones.

Análisis DAFO											
VERIFICAR			EXPORTAR								
Fecha de Inicio	Empleado	Factor Combinado	Tipo Riesgo	Estrategia	Valor	Cuenta de Cliente	Diferencia en	Diferencia en	Fecha de Implantación	Fecha de Finalización	Observaciones
16/01/2015	JUAN CUBERO	Riesgo	Fuerte Competencia	Descuento Adicional	5%	225	-25.196,63 €	-15.000,00 €	01/06/2015	31/09/2015	Estrategia temporal
16/01/2015	JUAN CUBERO	Riesgo	Cliente solo campaña	Oferta Familia Nueva	Retail	365	-17.491,56 €	1.500,00 €	01/07/2015	31/12/2015	Apoyo Marketing
16/01/2015	JUAN CUBERO	Limitación	Calidad	Compras Familia Obsoles		4.125,00 €	-4.218,88 €	0,00 €	01/06/2015	31/12/2015	se necesita Caducidad superior a 12 meses

Figura 42: Exportación Toma de Decisiones

En este punto el fichero se importa en el ERP y se exporta al sistema de BI a un nuevo esquema estrella que permite valorar grado de éxito de las estrategias, agrupando por empleado y estado.

5.5.5 Análisis de resultados

El objetivo del sistema determina que se tiene que poder contrastar el resultado de las estrategias implantadas, para tal efecto se diseña un informe funcional agrupado por fecha, Comercial, Estado de la Estrategia, Estrategia y Cliente. El comercial puede determinar el grado de cumplimiento de la Estrategia mediante el % Recuperado que indica la diferencia entre % Objetivo y % de Caída Original, % Días completados (sólo aplicable a estado Abierto, en Estrategias completadas este será del 100%) y finalmente el progreso hasta completar el objetivo marcado, el cual será inamovible en Estrategias completadas, determinando el éxito de la estrategia en caso de igualar o superar el 100%. En la Figura 43 se puede ver el resultado de una estrategia implantada.

RESULTADO ESTRATEGIAS CLIENTES						
Clientes	Diferencia Inicial	% Caída Original	% Objetivo	% Recuperado	% Días Completados	% Progreso
2014				7,50 %	11,11%	37,50 %
JUAN CUBERO				7,50 %	11,11%	37,50 %
Abierto				7,50 %	11,11%	37,50 %
Descuento Adicional				7,50 %	11,11%	37,50 %
SABATE DOLÇ I SALAT S.L.				7,50 %	11,11%	37,50 %

Figura 43: Resultado de Estrategias en Clientes

6. Conclusiones

En base a la asignatura Inteligencia de Negocio y Gestión de Procesos, el proyecto y la experiencia de Juan Carlos Trujillo y Alejandro Mate, mi percepción y conocimiento de las tecnologías, metodologías, modelos, negocio relacionadas con el Business intelligence, quedan patentes en este estudio.

Además de las soluciones expuestas en el proyecto, se ha completado la implementación del caso práctico en una organización, la cual ha requerido todo tipo de habilidades gerenciales asociadas generalmente con el liderazgo, coaching [\[33\]](#) o entrenar, inteligencia emocional o ingeniería social entre otros, el objetivo del proyecto no deriva solo en la generación de un modelo analítico y su explotación. Este proyecto tiene consecuencias directas en como la personas redescubren la información, pasando de ser consumidores a constructores. Este cambio repercute directamente sobre el usuario, el cual se transforma, consiguiendo un perfil más analítico descubriendo y comprendiendo los datos.

En el objeto del proyecto, la retroalimentación de estos sistemas permite crear inteligencia colectiva de las decisiones adoptadas, permitiendo compartir las decisiones exitosas y aquellas que fracasaron. Aportando la base de conocimiento de la estrategia de la empresa y una metodología concreta para el estudio de los Datos.

Es por todo esto, que hemos valorar este tipo de proyectos con cautela, ya que si bien esta tecnología es madura, hay gran cantidad de herramientas no pueden adaptarse rápidamente a los requisitos y objetivos del proyecto.

La retro alimentación ha de considerarse como un sistema evolutivo, ya que los factores internos y externos (factores DAFO), cambian continuamente. Se hace necesario el restudio continuo de estos factores, que temporalmente han de guardar relación con el plan Estratégico de la Empresa.

Tradicionalmente el proceso de toma de decisiones basado en BI, finalizaba con la decisión, el proyecto aporta una nueva fuente de análisis, que consiste en analizar el proceso de toma de decisiones desde el propio sistema de BI.

El proyecto tiene alta dependencias del usuario de negocio, el cual ha de adoptar una metodología de análisis concreta, la cual ha de adoptarse de forma disciplinada, ya que en otro caso la calidad de la información obtenida mediante la retro alimentación, se ve seriamente comprometida. Esta es sin duda una de sus principales limitaciones.

En general cuando se habla de BI, siempre hay alguien que utiliza la frase “Las mentiras de BI”, esta frase alcanza su comprensión en los procesos de análisis de los datos, ya que por norma general no se analizan los datos de manera disciplinada por todos los usuarios de negocio, llevando generalmente a conclusiones de la información incorrectas o sesgadas. En mi proyecto el usuario de negocio hace uso de metodologías DAFO, para alcanzar esta disciplina. Permitiendo convertir las decisiones de negocio, en nuevas transacciones para el análisis. Generando una base corporativa de decisiones analizadas y verificadas.

7. Referencias

- [1] [BI](#) Business Intelligence (Consultada por última vez el 08/04/2015)
- [2] [Data Discovery](#) Descubrimiento a través de los Datos (Consultada por última vez el 08/04/2015)
- [3] [DAFO](#) Metodología de estudio de la situación de una empresa o un proyecto (Consultada por última vez el 08/04/2015)
- [4] [Edgar Frank Codd](#) Científico informático inglés conocido por sus aportes a la teoría de bases de datos relacionales (Consultada por última vez el 09/01/2015)
- [5] [ERP](#) Sistema de planificación de recursos empresariales (Consultada por última vez el 09/01/2015)
- [6] [OLTP](#) Procesamiento de Transacciones En Línea (Consultada por última vez el 09/01/2015)
- [7] [CRM](#) Administración de la relación con los clientes (Consultada por última vez el 09/01/2015)
- [8] [Microsoft Dynamics AX](#) ERP de Microsoft (Consultada por última vez el 05/03/2015)
- [9] [William Kent](#) fundador y primer presidente del comité ANSI X3H7 en Gestión de la Información de objetos (Consultada por última vez el 02/03/2015)
- [10] [A Simple Guide to Five Normal Forms in Relational Database Theory](#) Guía creada por William Kent (Consultada por última vez el 02/03/2015)
- [11] [Desarrollos ad-hoc](#) Solución específicamente elaborada para un problema o fin preciso (Consultada por última vez el 02/03/2015)

- [12] [Almacén de datos corporativo](#) Data WareHouse, Colección de datos orientada a un determinado ámbito, integrado, no volátil y variable en el tiempo, que ayuda a la toma de decisiones (Consultada por última vez el 08/04/2015)
- [13] [Bill Inmon](#) Considerado el padre del Data WareHouse (Consultada por última vez el 08/04/2015)
- [14] [Ralph Kimball](#) Él es uno de los diseñadores originales de almacenamiento de datos (Consultada por última vez el 05/03/2015)
- [15] [BI clásico](#): Real-time Business Intelligence, BI Dirigido por los datos (Consultada por última vez el 02/02/2015)
- [16] Diseño y Explotación de almacenes de datos. Juan Carlos Trujillo, José Norberto Mazón y Jesús Pardillo. Editorial Club Universitario (Consultada por última vez el 06/12/2014)
- [17] [ETL](#) Extracción, Transformación y Carga (Consultada por última vez el 08/04/2015)
- [18] [OLAP](#) Procesamiento analítico en línea (Consultada por última vez el 05/03/2015)
- [19] [CMI](#) Cuadro de mandos Integral /Balanced Scorecard / DashBoard (Consultada por última vez el 05/03/2015)
- [20] [Data Mart](#) Es una versión especial de almacén de datos (Data WareHouse). Son subconjuntos de datos con el propósito de ayudar a que un área específica dentro del negocio (Consultada por última vez el 05/03/2015)
- [21] [Tipos de modelamiento de un DW](#) Esquema Estrella, Copo de Nieve y Constelaciones (Consultada por última vez el 01/02/2015)

- [22] [MOLAP, ROLAP, HOLAP](#) Soluciones principales utilizadas en Inteligencia de negocios que agilizan la consulta de grandes cantidades de datos (Consultada por última vez el 01/02/2015)
- [23] [KPI](#) Indicador clave de desempeño (Consultada por última vez el 01/02/2015)
- [24] [Cognición](#) facultad para procesar información (Consultada por última vez el 03/04/2015)
- [25] [Gartner](#) Empresa consultora especializada en Tecnología Informática (Consultada por última vez el 02/04/2015)
- [26] [Explicación del Cuadrante de comparativas Gartner](#) (Consultada por última vez el 05/03/2015)
- [27] [Magic Quadrant for Business Intelligence and Analytics Platforms](#) (Consultada por última vez el 05/03/2015)
- [28] [Microsoft SQL Server 2014](#) Principal servidor de bases de datos de Microsoft (Consultada por última vez el 12/04/2015)
- [29] [Microsoft SQL Server Data Tools - Business Intelligence para Visual Studio 2013](#) Entorno integrado para que los desarrolladores de software de bases de datos (Consultada por última vez el 12/04/2015)
- [30] [Directorio Activo](#) El Servicio de Directorio Activo proporciona la capacidad de establecer un único inicio de sesión en sistemas Microsoft y un repositorio central de información para toda su infraestructura, lo que simplifica ampliamente la administración de usuarios y equipos (Consultada por última vez el 08/01/2015)

[31] [Data Mining](#) La minería de datos o exploración de datos, orientado al proceso que intenta descubrir patrones en grandes volúmenes de conjuntos de datos (Consultada por última vez el 02/03/2015)

[32] [EDI](#): Intercambio electrónico de datos (Consultada por última vez el 01/12/2014)

[33] [Coaching](#) Método que consiste en acompañar, instruir y entrenar a una persona o a un grupo de ellas, con el objetivo de conseguir alguna meta o de desarrollar habilidades específicas (Consultada por última vez el 13/04/2015)